

DOCUMENT RESUME

ED 361 236

SO 022 621

AUTHOR Schrettenbrunner, H., Ed.; Westrhenen, J. van, Ed.
 TITLE Empirical Research and Geography Teaching.
 Nederlandse Geografische Studies 142 (Netherlands
 Geographical Studies 142).
 INSTITUTION Free Univ. of Amsterdam (Netherlands). Center for
 Geographic Education.; Royal Dutch Geographical
 Society, Amsterdam (Netherlands).
 REPORT NO ISBN-90-6809-152-2; ISSN-0169-4839
 PUB DATE 92
 NOTE 198p.
 AVAILABLE FROM KNAG/Netherlands Geographical Studies, P.O. Box
 80123, 3508 TC, Utrecht, The Netherlands.
 PUB TYPE Reports - General (140) -- Collected Works - General
 (020)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
 DESCRIPTORS Computer Uses in Education; *Educational Research;
 Elementary Secondary Education; Foreign Countries;
 *Geographic Concepts; *Geography Instruction; Higher
 Education; Map Skills

ABSTRACT

This volume contains 10 articles that concern research efforts undertaken in the field of geography education and training. Written by researchers from around the world, the articles include: (1) "Empirical Didactics of Geography: History of a Working Group" (H. Schrettenbrunner); (2) "The Development and Evaluation of a Geography Computer Simulation" (H. Schrettenbrunner); (3) "Some Psychological Guidelines for Designing Instructional Computer Simulations" (D. Leutner; B. Schreiber); (4) "Computer-Assisted Evaluation of Curriculum Software" (G. Hess); (5) "The Cognitive Complexities of Accessing Curriculum Specific Databases" (R. Gerber; J. Lidstone); (6) "Geographical Procedural Knowledge and Map Skills (J. van der Schee and others); (7) "How Do Students in Higher Education Perceive Maps?" (R. Gerber); (8) "Group Learning in Map Work" (B. L. Yau and others); (9) "Effectiveness of Gaming Method in Teaching Grid References and Conventional Signs: An Experimental Study" (Y. Fung); and (10) "Assessment in Geography: An Evaluation of the 'Solo' Taxonomy" (P. Stimpson). (DB)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

**H. Schrettenbrunner &
J. van Westrhenen**
Editors

Empirical Research and Geography Teaching

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL
IN OTHER THAN PAPER COPY HAS BEEN
GRANTED BY

J. G. BORCHERT

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

So 022 621

BEST COPY AVAILABLE

**Empirical Research
and Geography
Teaching**

Nederlandse Geografische Studies / Netherlands Geographical Studies

Redactie / Editorial Board

Prof. Dr. J.M.M. van Amersfoort
Dr. H.J.A. Berendsen
Drs. J.G. Borchert
Prof. Dr. A.O. Kouwenhoven
Prof. Dr. H. Scholten
Drs. P. Sijsma

Plaatselijke Redakteuren / Associate Editors

Drs. J.G. Borchert,
Faculteit der Ruimtelijke Wetenschappen Rijksuniversiteit Utrecht
Dr. D.H. Drenth,
Faculteit Beleidswetenschappen Katholieke Universiteit Nijmegen
Dr. A.C.M. Jansen,
Economisch-Geografisch Instituut Universiteit van Amsterdam
Drs. F.J.P.M. Kwaad,
Fysisch-Geografisch en Bodemkundig Laboratorium Universiteit van Amsterdam
Drs. P. Sijsma
Faculteit der Ruimtelijke Wetenschappen Rijksuniversiteit Groningen
Dr. L. van der Laan
Economisch-Geografisch Instituut Erasmus Universiteit Rotterdam
Dr. J.A. van der Schee,
Centrum voor Educatieve Geografie Vrije Universiteit Amsterdam
Drs. F. Thissen,
Instituut voor Sociale Geografie Universiteit van Amsterdam

Redactie-Adviseurs / Editorial Advisory Board

Dr. G.J. Ashworth, Dr. P.G.E.F. Augustinus, Prof. Dr. G.J. Borger,
Prof. Dr. J. Buursink, Prof. Dr. K. Bouwer, Dr. C. Cortie, Dr. J. Floor,
Drs. J.D.H. Harten, Prof. Dr. G.A. Hoekveld, Dr. A.C. Imeson, Dr. A.C.M. Jansen,
Prof. Dr. J.M.G. Kleinpenning, Prof. Dr. F.J. Ormeling, Dr. H.F.L. Ottens,
Dr. H. Reitsma, Dr. H.Th. Riezebos, Drs. P. Schat, Drs. F. Schuurmans, Dr. J. Sevink,
Dr. W.F. Slegers, T.Z. Smit, Drs. P.J.M. van Steen, Dr. J.J.J. Sterkenburg,
Drs. H.A.W. van Vianen, Dr. J. van Weesep

ISSN 0169-4839

Nederlandse Geografische Studies 142

Empirical Research and Geography Teaching

Edited by
H. Schrettenbrunner & J. van Westrhenen

Utrecht/Amsterdam 1992

Koninklijk Nederlands Aardrijkskundig Genootschap/
Centrum voor Educatieve Geografie Vrije Universiteit Amsterdam

Translation chapters 1, 2, 3, 4 and 6: J. Mulder
Copy-editing: H.L. Bijl

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Empirical

Empirical research and geography teaching / ed. by H. Schrettenbrunner & J. van
Westrhenen. - Utrecht : Koninklijk Nederlands Aardrikskundig Genootschap ; Amsterdam
Centrum voor Educatieve Geografie, Vrije Universiteit Amsterdam. - III. - (Nederlandse
geografische studies, ISSN 0169-4839 ; 142)

Met lit. opg.

ISBN 90-6809-152-2

Trefw. : leerplanontwikkeling / geografie-onderwijs.

ISBN 90-6809-152-2

Copyright © Centrum voor Educatieve Geografie Vrije Universiteit Amsterdam
The Netherlands 1992

All rights reserved. No part of this publication may be reproduced in any form, by print or
photoprint, microfilm or any other means, without written permission from the publishers.

Printed in the Netherlands by Elinkwijk b.v., Utrecht.

Contents

List of figures	v
List of tables	vi
Preface	1
1	EMPIRICAL DIDACTICS OF GEOGRAPHY - HISTORY OF A WORKING GROUP
	3
1.1	Introduction
	3
1.2	1984: A challenge for more awareness of methodology
	3
1.2.1	Deficiencies in the application of theory
	3
1.2.2	Deficiencies in the design of research work
	4
1.2.3	Deficiencies in the nature of variables
	5
1.2.4	Deficiencies in the use of statistical methods
	5
1.3	1986: Research perspectives for the empirical didactics of geography
	7
1.3.1	Macrotheories for a theory of teaching
	7
1.3.2	How profound should theory be?
	9
1.3.3	The design of experimental research
	10
1.4	1988: A memorandum for a working group
	12
1.4.1	Major guidelines for future research
	12
1.4.2	How to plan an experiment
	13
1.5	1990: Results of the working group on empirical research
	14
1.6	Conclusion
	16
References	17
2	THE DEVELOPMENT AND EVALUATION OF A GEOGRAPHY COMPUTER SIMULATION
	19
2.1	Basic assumptions for the development
	19
2.2	"Hunger in the Sahel", a simulation
	21
2.3	Strategies in the simulation
	23
2.4	Experimental evaluation of different versions
	25
2.4.1	The implementation of teaching functions (Leutner 1989)
	25
2.4.2	Checking the function of the adaptive aids (Leutner 1991)
	29
2.4.3	Optimizing the non adaptive aids (Leutner 1991)
	29
2.4.4	Controlling the effects in different school types
	33
2.4.5	Installing different levels of difficulty
	37
References	44

3	SOME PSYCHOLOGICAL GUIDELINES FOR DESIGNING INSTRUCTIONAL COMPUTER SIMULATIONS	47
3.1	Introduction	47
3.1.1	Background information on screen versus on paper	48
3.1.2	Inductive thinking, learning aids, and learning outcome	49
3.2	Hypotheses	50
3.3	Methodological design	50
3.4	Results	52
3.5	Summary and discussion	57
3.6	Conclusion	58
	References	59
4	COMPUTER-ASSISTED EVALUATION OF CURRICULUM SOFTWARE	61
4.1	Introduction	61
4.2	State of the art	61
4.3	The concept of the overall program package	62
4.3.1	Creation of test with the supereva program	64
4.3.2	The Eva program for use by testee	68
4.3.3	Hardware specific limitations	69
4.3.4	Problems between hardware and software	69
4.4	Initial observations on user interfaces and content design	70
4.4.1	User interface for the Supereva program	70
4.4.2	The user interface for the Eva test module	70
4.4.3	Design of test module contents	71
4.5	Initial results from evaluations with Supereva	71
4.5.1	Comparison of results between paper test and computer controlled test	71
4.5.2	Initial results of statistical evaluation	72
4.6	Conclusions	73
	References	73
5	THE COGNITIVE COMPLEXITIES OF ACCESSING CURRICULUM SPECIFIC DATABASES	75
5.1	Introduction	75
5.2	Factors that influence students' success with computer-assisted learning	75
5.2.1	Student characteristics	76
5.2.2	Learning environment	76
5.2.3	Computer assisted learning	77
5.3	Types of user interface found in curriculum specific databases	78
5.4	The empirical study	80
5.4.1	Method	80
5.4.2	Instruments	80

5.5	Analysis of the data	82
5.6	Results	82
5.7	Discussion and conclusion	84
References		85
6	GEOGRAPHICAL PROCEDURAL KNOWLEDGE AND MAP SKILLS	91
6.1	Introduction	91
6.2	Procedural knowledge	91
6.3	Research questions	94
6.4	The research method	94
6.5	The research population and time planning	96
6.6	Exploratory research	97
6.7	A few final remarks	111
References		112
7	HOW DO STUDENTS IN HIGHER EDUCATION PERCEIVE MAPS?	115
7.1	The problem	115
7.2	Design of the study	115
7.3	Instruments used in the study	116
7.4	The participants in the study	124
7.5	Analyses of the data	124
7.6	Point symbols	124
7.7	Linear symbols	126
7.8	Areal symbols	128
7.9	Combination of symbols	130
7.10	Discussion	131
References		133
8	GROUP LEARNING IN MAP WORK	135
8.1	Introduction	135
8.2	The mastery learning strategy	135
8.3	The cooperative learning model	137
8.4	The learning team model	138
8.5	The group learning strategy	139
8.6	The investigation	140
8.6.1	The sample	140
8.6.2	Method of study	140
8.6.3	Procedure	141
8.6.4	Results and discussion	141
References		143

9	EFFECTIVENESS OF GAMING METHOD IN TEACHING GRID REFERENCES AND CONVENTIONAL SIGNS: AN EXPERIMENTAL STUDY	145
9.1	Introduction	145
9.2	Methodology	146
9.3	Result	148
9.4	Discussion	149
9.5	Conclusion	150
	References	151
10	ASSESSMENT IN GEOGRAPHY: AN EVALUATION OF THE SOLO TAXONOMY	157
10.1	Introduction	157
10.2	The need for theory	157
10.3	The SOLO model	159
10.4	SOLO based ordered outcome questions in geography	160
10.4.1	An example of an ordered outcome question	161
10.4.2	The stem	161
10.4.3	Unistructural (U) and multistructural (M) level questions	163
10.4.4	Relational (R) level questions	164
10.4.5	Extended abstract (E) level questions	164
10.5	The validity of a SOLO format in geography	165
10.5.1	Experimental design	165
10.5.2	Results	167
10.6	Conclusion	174
	References	175
	List of contributors	179

List of figures

Figure 1.1	Formal structure of a theory of teaching	7
Figure 1.2	Two, three or six macrotheories for a theory of teaching: levels of organizational units: classroom, school, education	8
Figure 2.1	Rainfall and drought in the Sahel	22
Figure 2.2	Main menu of the "Hunger in the Sahel" simulation	23
Figure 2.3	Cognitive test: experimental conditions (adaptive vs. tutorial)	27
Figure 2.4	Action test: experimental conditions (adaptive vs. tutorial)	28
Figure 2.5	Cost-benefit pages	31
Figure 2.6	Correct answers in the retention test	32
Figure 2.7	Cognitive tests, raw scores, means per school type (experiment 1991)	35
Figure 2.8	Action test, years achieved, means per school type, level, tutorial (experiment 1991)	36
Figure 2.9	Cognitive test, raw scores, ATI-effects between school type and level of difficulty	40
Figure 2.10	Experienced difficulty in experiment 1991 (grade 6) and 1989 (grade 8)	41
Figure 2.11	Factor 1 "Interest" by difficulty	41
Figure 2.12	Students' attitudes towards the simulation (experiments 1989, 1991)	43
Figure 3.1	Mean posttest knowledge	54
Figure 3.2	Differential regression of posttest knowledge on inductive reasoning	55
Figure 3.3	Differential regression of posttest knowledge on inductive reasoning	55
Figure 4.1	Supereva/Eva overall system	63
Figure 4.2	Screen display of a learning goal oriented achievement test	64
Figure 4.3	Eva testee program - question in a personality test	67
Figure 4.4	Eva testee program - on-line Help window	68
Figure 4.5	Item analysis - graphics display (here pie-chart)	70
Figure 6.1	Population distribution in Alpes du Nord	93
Figure 6.2	Research design for research question 2	95
Figure 6.3	Question 1 about Finland from the map skills test	98
Figure 6.4	Question 2 about France from the map skills test	99
Figure 6.5	Question 3 about Bergezië and Geria from the map skills test	100, 101
Figure 6.6	Question 4 about Bordeille from the map skills test	102, 103
Figure 6.7	Question 5 about Vasporije from the map skills test	104
Figure 6.8	Categories for student reactions to map skills tests	110
Figure 7.1	Map 1 - Settlements in Bentinck	118
Figure 7.2	Map 2 - Settlements in Comorin	118
Figure 7.3	Map 3 - Settlements in Dargaville	119
Figure 7.4	Map 4 - Settlements in Eaglehawk	119
Figure 7.5	Map 5 - Road network in Ferrara	120
Figure 7.6	Map 6 - Road network in Good Hope	120
Figure 7.7	Map 7 - Road network in Heathecote	121
Figure 7.8	Map 8 - Road network in Kandrian	121
Figure 7.9	Map 9 - Population density in Lethbridge	122
Figure 7.10	Map 10 - Population density in Mostar	122
Figure 7.11	Map 11 - Population density in Parinas	123
Figure 7.12	Map 12 - Population density in Sherpa	123
Figure 8.1	Teaching processes of the mastery learning strategy	136
Figure 8.2	Teaching processes of learning group strategy	140

Figure 10.1	A typical multiple choice geography question	158
Figure 10.2	An example of a SOLO superitem in geography	162
Figure 10.3	SOLO superitems used in the analysis of data	171
Figure 10.4	Land use in Hong Kong (1987)	172

List of tables

Table 2.1	Multivariate analysis of variance	32
Table 2.2	Repeated measurement analysis of variance: cognitive tests per school type, level, tutorial, within-subjects effects	35
Table 2.3	Analysis of variance: action test by treatment	37
Table 2.4	Repeated measurement analysis of variance, cognitive test, school types and difficulty (experiment 1991), between-subjects	40
Table 2.5	Factor analysis, factor matrix, principal components analysis, unrotated	42
Table 3.1	Versions of the computer simulation "Hunger in the Sahel" and no. of subjects N within each cell of the experimental design	52
Table 3.2	Analysis of covariance (ancova) of posttest knowledge	53
Table 3.3	Means and standard deviations of cognitive variables	56
Table 3.4	Pearson correlations between cognitive aptitudes and posttest knowledge	56
Table 4.1	Scores on paper test and computer test	72
Table 5.1	Correlation analyses on success and efficiency ratings for selected geography databases using Spearman's correlation co-efficient (all students)	86
Table 5.2	Correlation analyses on success and efficiency ratings for selected geography databases using Spearman's correlation co-efficient (year 8)	87
Table 5.3	Correlation analyses on success and efficiency ratings for selected geography databases using Spearman's correlation co-efficient (year 12)	88
Table 5.4	Correlation analyses between success and efficiency ratings for selected geography databases and selected independent variables using Spearman's correlation co-efficient	89, 90
Table 6.1	The percentage of right and wrong answers per question to the four variants of the map skills test	106
Table 6.2	The percentage of right answers in the open and closed versions of the map skills test as compared to the total number of right answers per question	107
Table 6.3	The relationship between foreign labor and average income on the map of the first question in the map skills test	108
Table 6.4	Percentage distribution of answers to question 2A for the open versions of the map skills test	109
Table 7.1	Block design used in visual hierarchies study	116
Table 7.2	Students' preferences and understanding of visual hierarchies on thematic maps using point symbols	125
Table 7.3	Frequencies of students' reasons for identifying different size of settlements on maps 1-4	125
Table 7.4	Frequencies of students' selection of the most visible symbols on maps 1-4	125

Table 7.5	Spearman's correlation co-efficients between students' perceptions and understandings of visual hierarchies on maps 1-4	126
Table 7.6	Students' preferences and understanding of visual hierarchies on thematic maps using linear symbols	127
Table 7.7	Frequencies of students' reasons for identifying the level of importance on roads on maps 5-8	127
Table 7.8	Frequencies of students' selection of the most visible symbols on maps 5-8	128
Table 7.9	Spearman's correlation co-efficients between students' perceptions and understandings of visual hierarchies on maps 5-8	128
Table 7.10	Students' preferences and understanding of visual hierarchies on thematic maps using areal symbols	129
Table 7.11	Frequencies of students' reasons for identifying the populations of settlements and the different population densities of areas on maps 9-12	129
Table 7.12	Frequencies of students' selection of the most visible symbols on maps 9-12	130
Table 7.13	Spearman's correlation co-efficients between students' perceptions and understandings of visual hierarchies on maps 9-12	130
Table 7.14	Selected combination of map symbols from maps 1-12	131
Table 7.15	Reasons for the selection of the combination of map symbols from maps 1-12	131
Table 8.1	A comparison of scores between the experimental and control groups	142
Table 9.1	Learning outcomes of two experimental groups and one control group	149
Table 10.1	Results of the partial credit modelling	168
Table 10.2	Item quality	169
Table 10.3	Percentages of pupils answering each subitem correctly	170

Preface

Keeping up with the "state of the art" in the field of geographical education research is one of the most important tasks of the Commission on Geographical Education of the International Geographical Union.

Norman Graves published a first overview eight years ago when he edited a collection of contributions in his book "Research and Research Methods in Geographical Education" (1984). The different chapters contained reports on historical, qualitative, hermeneutical- and quantitative-empirical research.

Helmut Schrettenbrunner issued a challenge to the audience of the Freiburg Symposium in 1984 when he asked for more theory-based, hypotheses-oriented, experimental, quantitative empirical research. His goal was to start a broad discussion on an international agreement about scientific standards for research in geographical education.

The Commission is pleased to note that Helmut Schrettenbrunner has succeeded in assembling a Working Group on Empirical Research and we are sure that this publication will be a milestone for our research philosophy.

This book contains many criteria for designing and evaluating future research projects. Some contributions in this book focus on the development and testing of geographical software. Other contributions focus on maps, teaching modes and testing in geography teaching. The different authors present arguments for considering empirical didactics of geography as an applied science.

The Commission congratulates Helmut Schrettenbrunner and his colleagues for these impressive research activities and hopes that the readers will appreciate this book.

Hartwig Haubrich

(Prof.dr. Hartwig Haubrich is chairman of the IGU Commission Geographical Education)

1 EMPIRICAL DIDACTICS OF GEOGRAPHY - HISTORY OF A WORKING GROUP

Helmut Schrettenbrunner

1.1 Introduction

Books are usually regarded as the solid result of an author's endeavor and may be accepted as "a written authority". In our case, the book is also a document of an eight year migratory passage to the final result. It all started in Freiburg 1984 (Germany), then went on to Sitges 1986 (Spain), to Brisbane 1988 (Australia) and Hong Kong 1990, always with Boulder (USA) as the final target in the mind.

This book on empirical research shows how the working group within the IGU commission "Geographical Education" formulated their aims, which are documented by the author's respective contributions to the above named IGU congresses.

This article is a condensed version of those aims and demonstrates how long it can take to implant an idea in an IGU commission and to bring it to a published result. The reader is kindly asked to consider the first grumbling comments on the "deficiencies" found in one of the congress publications, then to note the first "postulates" for improved quality and last but not least to read the first examples of empirical case studies.

This article documents the long and painstaking process of striving for an international standard of research, with contributions from Australia, Hong Kong, the Netherlands and Germany.

1.2 1984: A challenge for more awareness of methodology

1.2.1 Deficiencies in the application of theory

If empirical methods and quantitative analyses are to make sense they must be based on theoretical assumptions which must have such a degree of accuracy that they can be formulated in a set of hypotheses. One would imagine that this goes without saying, because otherwise we would either have theoretical speculations without concrete proof (no detailed and operable hypotheses, no empirical application and testing) or the mere counting of flies' legs with little sense.

Let us add a few fictitious examples in connection with our theme "media and their evaluation", the title of the Freiburg Symposium in 1984.

We have interesting resumes of theories from education or psychology, such as information theory, perception theory, and cybernetical theory, which describe

possible approaches to parts of a very complex phenomenon, i.e. how can we interpret the interactions between medium, teacher, pupil, class, objectives and geographical content?

It is very worthwhile to keep the complex situation of using media in school in mind, but generally speaking the result of this type of resumé is that we formulate problems which are too complex for empirical research and are unable to control the complex situation in school (teacher effects, group effects, situation effects of media). In short, the scope of our research is often too vast, our hypotheses, in consequence, are inoperable and statistical methods of acknowledged standards cannot be applied.

Postulate 1: It is preferable to have a specific and limited theoretical basis (set of hypotheses) to control side effects and apply standard methods of statistics, rather than to deal with the whole complex situation of media in school with no chance of empirical testing.

1.2.2 Deficiencies in the design of research work

If the findings of our empirical data of teacher and pupil interactions with media are to make sense, the methods of investigation must be chosen carefully.

This is not so much a question of how many people are investigated, as which aims we set and what types of interactions and hypotheses we choose.

To make one point clear right from the beginning, there is no analysis with the statistical quality of representativity in the field of education in geography, even if most research work seems to imply this in titles, e.g. "The perception of Europe by Italian students", "Children's world perception", "Activities prompted by geography books". These titles suggest a methodological quality which they cannot maintain because the concept of representativity does not consist of a certain minimum number of cases, rather essentially of a theoretical idea of selection.

Other sciences such as demography, opinion research and, again, psychology can render good examples for representative surveys, but researchers of education in geography have not yet taken note of them. As a result they add case after case to a meaningless set of data. So what can we do if representativity has not yet been reached and is very difficult to obtain for technical and organizational reasons?

There is one realistic solution: We have to change our aims and questions and choose experimental designs. Representative descriptive surveys give us description of all attributes of a defined sample (e.g. all German students of geography) and answer the question, "how" certain attributes are distributed.

In contrast, experimental design does not give descriptions of the independent variables for a defined sample, but enables us to test causal relations between variables and answer the question, "why" certain connections exist (e.g. "why do male students react less intensively to visual representation?").

This kind of experimental design can also be processed with adequate statistical methods (e.g. analysis of variance) and can help us more than a simple description ("how do pupils react to textbooks?", "How is textbook revision being done?", "What is the world perception in India's education?"), because it gives explanations ("Why are type A maps more difficult for 11 year old children than type B?", "Why is video-presentation of complex cognitive structures less effective than teacher-presentation?").

Postulate 2: *Because representativity is practically non-feasible, empirical analyses should have quota bases and experimental character, they should place less emphasis on the description of samples and more emphasis on the explanations of independent and dependent variables.*

1.2.3 Deficiencies in the nature of variables

If our intended explanation of interactions is to be meaningful, our variables must be relevant and adequately complex. Let us consider the normal research work on the following independent variables: sex, age, mark in geography, experience with medium, attitude towards geography-teacher and in rare cases even pupil's IQ. Despite all of the considerations concerning theory (see 1.2.1) this reduced set of independent variables is used, because there are no other variables at hand.

Organizational and technical necessities reduce our theoretical framework to a fragment and lead to an explanation of variance far below 50%. This means that the majority of explanations needed in our context is beyond the scope of our analysis or that we have obtained only a minimum amount of explanations. And if we did not have a strong (but equally known) variable like age, these would be reduced to almost nothing.

Education is a very complex matter and needs adequate variables, for example; complex measures of intelligence, or a complex index for willingness to work, or a complex index for specific interests, and not just a single question for experience ("Did you work with films before?", "What mark did you receive on your last report?") or for interest ("Do you generally like geography?", "Do you prefer medium A or B?").

Such isolated question variables are of no value because they reduce any theory of education to non valid correlations. Therefore complex measures based on ample test batteries are necessary and have to be seen in a far wider range than only for geographical purposes. Test batteries to explain pupils' reactions in a learning program included subtests such as neuroticism, completion of sentences (language performance), attitude towards work (laziness and intellectual willingness), etc., which we term "ability to learn". Only then are we able to explain more than 50% of the variance, certainly not very much more because still too many variables are left outside our experimental design.

Postulate 3: *Because education is a very complex process, our didactic variables (domain specific knowledge; dependent, independent) must be based on complex test batteries.*

We have to take into account that results in one subject (e.g. geography) cannot be reduced to one-question-variables and cannot be seen in too narrow a context (the geography mark can be totally insignificant, general language performance may be of importance when working with geography media).

1.2.4 Deficiencies in the use of statistical methods

If the counting of events and attributes is to be meaningful, we have to apply adequate statistical methods unless we revert the meaning of quantitative research to its lowest level (i.e. just using figures instead of words).

Some of the most common errors are using statistical procedures in combination with inadequate data levels (e.g. wrong measurement of correlations), using only

bivariate correlations (e.g. mark in geography, interest in maps) where multivariable treatment should be applied and using procedures without control of necessary prerequisites (test of normal distribution, test of equal variance; correction through normalization). Generally speaking we can say that quantitative researchers in the field of education in geography have worked their way through the handbooks of statistics and reached a level which enables them to apply complex analysis (analysis of variance, factor analysis).

Statistical procedures and computer software packages, in turn, have then influenced the perspectives of research ("we research things that can be done with SPSS").

Undoubtedly, the methods used are solid and have a long and respected tradition. We have, however, not yet noticed that there are more modern methods which would fit our specific requirements better. Many of our variables are non-metric and therefore have always caused problems in multivariate analysis (Grizzle-Starmer-Koch-model, program NONMET), without recurring to highly complicated log-linear estimations.

Postulate 4: Because educational variables are often non-metric and designs multi-variate, we need adequate statistical methods and should apply non-metric path models, which suit any experimental design.

This paper is not going into detail with an analysis of the shortcomings of most of the research work done in our field. Some major defects can, however, be detected:

1. A nearly unlimited range of aims, themes and approaches, i.e. the total variety of subject matters reflecting education. It will be suggested that this range should be considerably limited.
2. A non-existing or haphazard use of theories with the subjective excuse that the researcher could not find a general theory of education and so might as well produce some sort of private theory himself. It will be suggested that a structured set of macrotheories can help us out of the maze, and again that concentration on a few fields is recommendable.
3. Deficiencies in the definition of terms and in the formulation of hypotheses, very common until now because of the unlimited range mentioned above, make solid empirical work impossible. Independent variables remain vague and sometimes also change position with dependent variables. It will be suggested that a multi-step procedure of operationalization should be applied.
4. Lack of appropriate data processing or misuse of statistical methods has been the result of the above mentioned deficiencies, which lead the researcher to ask at the end how he could polish his data with elegant statistical methods. It will be suggested that statistical methods should be considered not as a final embellishment but as necessities of theoretical considerations taken at the beginning of the research work.

1.3 1986: Research perspectives for the empirical didactics of geography

1.3.1 Macrotheories for a theory of teaching

Teaching theories suffer from common inadequacies: they deal with diverging aspects, treat different problems and are located at different levels. Our researcher is left alone in his choice of theories, and whichever he chooses it will not be comparable to others, so that his results will remain unique pieces in an imaginary mosaic which can never be completed because other research pieces will be found on other levels of abstraction or under different perspectives. "This situation could be perceived as discouraging if one looked for one theory of teaching that would be adequate for all instructional problems. Alternatively, it could be seen as reflecting the fact that these various attempts are more or less useful for different purposes, so that each of them possibly has its appropriate place in a larger frame of reference. Such a frame of reference can be provided by an all-encompassing theory of teaching if it is conceived as a hierarchical theory (figure 1.1). The subtheories may differ with respect to their levels of generality, the purposes they are developed for, or both aspects. Moreover, the lower a subtheory is located in the hierarchy, the closer it is to reality and research. A hierarchical consumption of a theory of teaching is advantageous in some respects: it enables one to integrate the various so far unrelated theories into a comprehensive pattern that reflects the different relationships between them. Moreover, the system as a whole need not be altered if a particular subtheory fails to be successful (Klauser 1985, p. 6)".

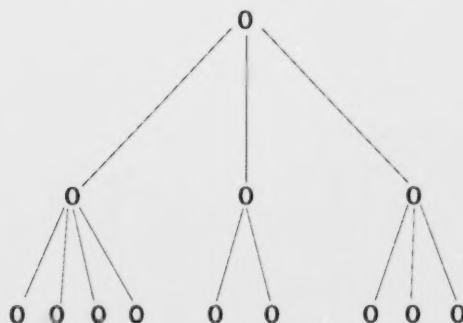


Figure 1.1 Formal structure of a theory of teaching

Klauser suggests a unifying theory of teaching which he presents on a two-dimensional table, one dimension standing for the type of question that the theory should be able to answer. He thereby makes three distinctions: descriptive, prescriptive and normative types of research. Educational psychology and empirical-analytical research on teaching are good examples for the description as well as the explanation of a section of reality. Normative studies deal with ultimate goals or general objectives and ethical standards, and should be considered under the label of

philosophy of education. The prescriptive type of science researches normatively established objectives with empirical analytical methods, thus forming a technology to solve problems of importance in practical school life. Corresponding to these distinctions, the theory of teaching can be conceived as consisting of two, three or six first-order subtheories or macrotheories (figure 1.2). The six cells depicted in the matrix need some comment.

Type of question	Type of study				
		Descriptive	Prescriptive	Normative	
	What?	A Objectives/ subject matter in classrooms	B Curriculum	C Ultimate ends	
	How?	D Teacher- student interactions	E Teaching methods	F Professional ethical standards	
		Educational psychology (Teaching research)	Educational technology or science of teaching design	Philosophy of education	

Figure 1.2 Two, three, or six macrotheories for a theory of teaching; levels of organizational units: classroom, school, education (from left to right) (Klauer 1985, pp. 6 - 7)

The discipline of Cell A is not yet well established, although it represents our domain specific knowledge. It is concerned with empirical research in objectives found in teaching processes. Disclosing the hidden curriculum, analyzing the content of a lecture, estimating the rate of repetition of a grammatical rule, or uncovering the flow of a piece of information in a lesson would be examples of research projects belonging to such a discipline. Cell B is concerned with curriculum theory. This discipline answers the question of what to teach, given that an ultimate goal has previously been established. The decision-making about such an ultimate end is achieved through the normative study of Cell C. Researchers have paid much attention to Cells D and E. The descriptive research in teacher-student interactions has been carried out not only by observational studies (Rosenshine & Furst 1973) but also by a large number of experimental investigations (Dunkin & Biddle 1974). The prescriptive study of teaching methods is precisely the type of research that Gage (1978) postulates because of its importance for improving educational practice. Finally, Cell F contains a normative discipline which is concerned with professional ethical standards for teachers (Klauer 1985, p. 7). When we draw our conclusions for the didactic research in geography we can delete

the normative type (philosophy) of research because it does not use empirical methods. The remaining four cells may be considered to fit empirical needs, with an emphasis on the theory of teaching methods.

Suggestion 1: Further research should center

1. mostly on teacher-student interactions and teaching methods in geography (theory of teaching methods, cells D and E),
2. partly on objective/subject matter in classrooms on curriculum theory (cell A and B), and not
3. on the philosophy of education (cell C and F). (Cell A on "objectives/subject matter in classrooms" which so far has not been sufficiently elaborated has been omitted)

These restrictions, hopefully, will exclude firstly all wide-angle and all-purpose efforts which admittedly have a purpose in the prevailing practical needs but cannot deal with the complex problems in a scientific way. Secondly these restrictions will exclude all mixture types of research across two or more cells. Further papers should examine the existing theories of teaching methods and give suggestions which to use and how to use them properly.

1.3.2 How profound should theory be?

If we take into consideration the results of empirical research in the didactics of geography we have to state that only a limited number of independent variables has been taken into account, such as age, grade, gender, marks, attitudes, IQ, performance in school and little more. This is partly due to the fact that these variables are at hand without greater problems (apart from IQ or attitudes) and also that we have never thought about an appropriate set of independent variables for our aims. If we try to explain, for example, the results in programmed instruction with the above mentioned variables, we are disappointed that less than 30% of the variance can be explained statistically. This should be evidence enough to show that our theoretical basis (expressed through the type of independent variables) is not profound enough to explain such a complex process as learning with a program. Other research which has taken into account a number of more complex characteristics (pupils' aptitude for guided work, pupils' willingness, expectation of success, fear of failure, power of concentration, motivation for speed, measured by test batteries, Rollett 1978) thus describing in a different form pupils' learning characteristics, can explain more of the variance (48%), because its theoretical background is of a different type and can therefore claim to be more valid.

Let us consider, however, that the above-named more complex pupils' characteristics are typical for a different science, i.e. for educational psychology and not for the didactics of geography. Our research aim is not so much the inner structure of a pupil, and if we are not to tangle with other research domains but are still interested in the psychological, theoretical explanation of the effects of geography teaching, then we should aim at joint research ventures, thus combining a psychological theory of learning with methods in or effects of geography teaching or learning. The solution should therefore be, if we strive for more complex theoretical explanation, to seek cooperation with neighboring sciences and not to trespass on unknown grounds.

Suggestion 2: Interdisciplinary cooperation must be sought if we want to widen the theoretical explanations of more complex relationships in geography teaching, or if we want to explain more about teachers' or pupils' behavior. Thereby we can include (with the cooperation of a colleague of a neighboring science) either another set of independent variables (not typical for the didactics of geography, such as personality variables) or another range of methods (such as Thematic Apperception Test, physiological reaction tests, etc.).

In cases other than interdisciplinary cooperation a very strict adherence to a much narrower theoretical basis should be followed, as will be shown in the next chapter.

1.3.3 The design of experimental research

It could already be shown that the concept of representativity, up to now accepted by our Commission (e.g. How do Danish pupils value other nations?), is wrong and has led to a number of meaningless publications.

It has equally been suggested that experimental research design should be applied instead, where R stands for the amount of reaction (dependent variable) which a person P shows when confronted with a number of stimuli S (independent variable). This mapping sentence can then be written:

P S $\Rightarrow$ R

We have to keep in mind the different types of operational definitions for the variables, which generally means for the dependent variable a set of definitions concerning measurements (number of seconds, sum of raw points, degree of attitude, etc.), whereas the independent variables are normally defined by prescriptions regarding the construction of the experiment (intensity of training, teacher experience yes/no, methods A/B etc). Nussbaum (1985, p. 64) points out that precise definitions are most important and shows how the use of the so-called facet design can be helpful.

"For a more detailed definition of our three variables P, S, R we introduce the so-called facets in the mapping sentence, which substructure our variables and give them a new order. The question of which facets shall be accepted within the mapping sentence is ultimately the researcher's decision, because we are dealing with definitions This method will now be demonstrated with one example drawing on the variable 'degree of tutoring by the teacher' (Nussbaum 1985, p. 65)".

Elements: **P S $\Rightarrow$ R**

Facets: **B A C D E F G**

Persons P will now be defined as pupils B with three different levels of learner's preconditions. Facets A C D E specify stimuli S, facet A indicating over which range our variable 'degree of tutoring shall be varied. 'A' is given on a continuous scale ranging from a very low to a very high degree of tutoring. In empirical research there should be a minimum of three experimental conditions which should be placed on points lying far away on this continuous scale to guarantee the eventual proof of

non-linear effects on the dependent variable. Facet C distinguishes only roughly between two types of information (content, learning aids). Facet D structures the content of learning by applying a similar concept in cognitive psychology. Facet E distinguishes 'work alone' from 'work in groups' because we can suppose that missing tutoring on the teacher's side can be compensated only by higher learning activities on the pupil's side. Facets F and G give intensities and different types of effects. If we now combine all possible facets (apart from A and F) we receive the following amount of structuples:

	P	S	R
1.	b1	c1 d1 e1	g1
2.	b2	c1 d1 e1	g1
.			
.			
.			
240	b3	c2 d4 e2	g5

"Each of these structuples represents a formalized definition of our independent variables S including a specification of our group of pupils P and the dependent R. The first structuple b1c1d1e1g1 could therefore be read as follows: The (low/high) intensity with which a teacher tutors his pupils, who have above average learning prerequisites, during their content learning of concepts (termini) in individual work results in (very low/very high) immediate learning effects of these pupils. Altogether we can use the mapping sentence as a system for definitions and extract a sum of 240 ($= 3 \cdot 2 \cdot 4 \cdot 2 \cdot 5$) sentences of this kind, which in their turn contain one definition after another of the structure under investigation" (Nussbaum 1985, pp. 68-69).

This example, which was drawn from Nussbaum's paper, should help us to visualize more vividly in which direction and by which means we must go. We have some similar approaches (Heilig 1982; 1984) and should be able to define a narrower and at the same time a more specific concept of what our independent variables within the didactics of geography should be. Experimental research without this overall definition of our field of interest is not feasible.

Suggestion 3: We should strictly adhere to the pledge to define our termini in a correct way and to define our variables in an operational way which can elucidate the amount of hypotheses within an experiment. We should also come to a conclusion within the commission on how to define the range of our research work by listing the relevant independent variables.

It should be remembered that statistical methods which are not being treated in this paper do not have a value of their own, but must fit adequately the chosen research design.

1.4 1988: A memorandum for a working group

Researchers in the empirical didactics of geography usually face some crucial problems:

1. What are specific characteristics of our research? - a question of distinction between other sciences (e.g. empirical pedagogic), see also figure 1.2, positions A, B, D, E.
2. What is our theoretical basis? - a question of hierarchical order of sciences (e.g. pedagogical and psychological theories explain students' behavior, also if applied to geographical content learning).
3. What does a standard empirical design consist of? - a question of methodological expertise (which our geographer may not have because his university education did not include the required methods).
4. What statistical methods should be applied? - in most cases a question put in the wrong place, i.e. after the collection of data and not together with our points 2 and 3.

Useful help with these problems cannot come from within the didactics of geography, but from outside, from sciences with a solid history of empirical research.

1.4.1 Major guidelines for future research

To avoid the above mentioned crucial problems we should accept the advice given by empirical educationalists (Leutner 1988a, p. 1, parentheses by H.S.): "The idea of empirical research in learning (with respect to geographical contents and objectives) is to formulate prescriptive sentences in order to optimize the effects of teaching (geography)". And when we outline a specific work of research we should think of the following prerequisites:

1. There must be a set of hypothetical effects based on theoretical grounds, e.g. "remedial instruction using various media (slides, films, abstract models) will enforce long-time retention effects because each of the media facilitate multi-dimensional memory structures".
2. Within our research we must analyze the variation of our treatments, e.g. "given a certain identical geographical content, measure the effects of using only slides versus using slides and abstract model together".
3. Our independent variables (small in number) must be of a kind that the geography teacher has a real chance to manipulate them, e.g. "the presentation of a given content in form of a black and white sketch or a four-color topographical map or a multi-colored, three-dimensional landscape drawing" (and not: variables such as gender, social background etc. of students; only when theory suggests possible differences should we investigate in such variables, e.g. map reading capabilities between boys and girls are significantly different as can be taken from research results on intelligence).
4. Our dependent variables (also small in number) must be such as to measure our hypothetical learning effects, e.g. "correct answers measured as raw points in a given item, or degree of performance measured as time consumed in producing a landscape profile from the basis of a topographical map".
5. Our research must consist of small experiments, e.g. "the variation of 2-4 methodological approaches to introduce map reading" (and not: the whole sum of possible differences of two complete curricula of geography).

1.4.2 How to plan an experiment

Leutner (1988a, p. 11) gives an illuminating example how to plan and organize an experiment (which for our specific purpose might be imagined to have geographical contents: an ecological model of agricultural land use in arid zones).

A. Theoretical background

When pupils work with dynamic systems on a computer we can assume that they learn details of the ecological situation presented (geography content) and of the structure of a model (methodological content). There remain two open questions:

- What role do those rules (algorithms, formulae, correlations) play which are hidden within the model and should be detected by the student? Here we have to recur to theoretical considerations on the learning of rules and on learning by discovery (Egan & Greeno 1973).
- What differences in learners' effects do we receive from the typical computer situation where 3-4 students work together compared to the normal learning situation? Here we have to recur to theoretical considerations concerning the principle of cooperative problem solving (Johnson 1981).
- Our attention must be drawn to the fact that neither theoretical background comes from geography but from pedagogic and from educational psychology (there is, of course, also a theory of geography: within the ecological model).

B. Mapping sentence

High/low learning effects can be explained by

- the high/low intensity
- in which learning contents are presented
- when students work alone/in groups
- and have to learn rules
- and can be divided in good/low pretest groups.

C. Experimental design

2 x 2 factorial design, analysis of covariance (i.e. $2 \times 2 \times 20 = 80$ students, as minimum of 20 per cell should be observed). Our independent variables (predictors for learners' effects) are

- information (did students receive information on the implicit rules of our model, yes/no),
- learning situation (did students work in groups or alone).

		information about rules	
		yes	no
learning situation	group	20	20
	alone	20	20

Pretest results and several parameters of steering the model were taken as covariates. Our dependent variable (criterion) consists of test results (a learning and retention test comprising all system rules, administered immediately after the computer simulation).

D. Possible results

This experiment could guarantee a solid answer to the methodological problem of how best to teach a geographical theory: What prerequisites must the teacher think of, how should he organize the learning situation? The results would give practical suggestions for real school situations and make the didactics of geography an applicable science.

Our example may have demonstrated that the didactics of geography is a subdiscipline of educational science. If we agree that our subject is the science of the institutionalized teaching and learning of geography we can define several major tasks:

1. Analyses of general teaching/learning issues with specific geographical contents with a certain danger of trying to invent the wheel again because educationalists, colleagues in general didactics or educational psychology might analyze the same problems only with different subject contents. But we should also see the great chance here for cooperation.
2. Analyses of the specific contents or methods of geography: Working with maps, atlases, globes, reliefs, models; investigations during outdoor situations, excursions, interviews; working with sketches, drawings of landscapes, of diagrams; data collection and processing and other techniques mostly and exclusively reserved for geography. The aim of our research in this case should be to find out how the specific contents and methods of geography should be taken into account to optimize learning them.
3. Analyses of detailed sets of geographical objectives with the intention of describing hierarchical trees of difficulties and giving indications of how and in which order best to proceed in teaching situations. We could then proceed and analyze transfer effects on certain hierarchical levels (within geography contents or also more general contents: e.g. understanding of complex models in other subjects).
4. Analyses of teaching methods which are the most helpful for our geographical contents and methods, e.g. types of remedial instruction in a computer simulation.

1.5 1990: Results of the working group on empirical research

There has been good contact between the active members of the working group since the 1988 congress in Brisbane, so that an overview can be given of the papers prepared for the IGU congress in 1992 (see the following chapters of this book).

2. Schrettenbrunner - The development and evaluation of a geography computer simulation.
3. Leutner & Schreiber - Some psychological guidelines for designing instructional computer simulations.

4. Hess - Computer-assisted evaluation of curriculum software.
5. Gerber & Lidstone - The cognitive complexities of accessing curriculum specific databases.
6. Van der Schee et al. - Geographical procedural knowledge and map skills.
7. Gerber - How do students in higher education perceive maps?
8. Yau et al. - Group learning in map work.
9. Fung - Effectiveness of gaming method in teaching grid references and conventional signs: an experimental study.
10. Stimpson - Assessment in geography: an evaluation of the Solo taxonomy.

Let us analyze some aspects of these contributions in order to find out if our endeavors to set a common basis and quality level has been successful.

Most of the authors have chosen the experimental design comparing different learning materials or treatments, which should be standard and are customary to find out the effects of different learning situations, different approaches or cartographic presentations. One author compares the results an experimental group has in a game with a control group, which also is a good way of finding out the effects of new methodological approaches.

And yet another author wants to find out if a specific taxonomy is of any value for geography. This is possible, but does not need an experimental group design.

It is important to note that the authors investigate variations of teaching methods, games or computer simulations and maps and conventional signs. This means that all of the authors have in mind research on topics that are relevant for future use in every-day school life.

This can be proved by a short glimpse at the dependent variables chosen in the experiments. All authors measure learning effects (posttest minus pretest scores, cognitive dimensions) and evaluate the effects of their materials, versions or groups. One author tries to go beyond simple quantitative measurements and emphasizes especially qualitative differences within a taxonomy, whereas the others implicitly use different levels without necessarily questioning the theoretical background. It is important to note that there is good reason for choosing learning effects with geographical material (such as maps, computer simulations, etc.) because the focus of an educational geography or didactics of geography should mainly be directed towards the issue of how to optimize geographical teaching material.

A closer look at the independent variables used in the papers is necessary because this proves what the authors believe to be the reasons for different learning effects and consider to be manageable by teachers or text-book writers.

Here we find:

- * variations in teaching strategies;
- * versions of material;
- * presentation of maps, symbols;
- all variables that could really be tackled by the classroom teacher and could lead him to better teach geography.

But there are also included other independent variables which explain effects through personality or intelligence dimensions, such as:

- * attitudes, motivation;
- * grade, age, gender;
- * self-esteem, anxiety (towards computers);
- * pupils' interactions, etc.

These variables indicate that also a minimum background of general psychological and educational theory serves as a basis.

And finally some remarks about statistical procedures employed. Group comparisons in 2 x 2 factorial designs using analysis of variance or covariance are named by most authors. And it seems that the efforts having been undertaken by the working group for years now will produce a promising contribution to the congress in 1992.

1.6 Conclusion

Our contribution and the articles for the book might help us to remember some of the important items an empirical analysis should consist of:

- a good theoretical basis (partly on a general psychological or educational background, partly on geographical and presentation modes);
- detailed hypotheses which must be manageable in the experiment (mostly explaining differences in learning effects with geographical material);
- themes of relevance for the teacher;
- an experimental design (usually testing different versions of material, or comparing different groups);
- an adequate statistical procedure to prove the hypotheses (mostly analyses of variance, covariance).

Further research was implied in the three suggestions:

1. survey of existing theories of teaching methods;
2. subject matters for interdisciplinary research;
3. definition of our independent variables and a point of view for future research.

Furthermore we should take into consideration a clearing board consisting of the commission's chairman and some other members who must be willing to act in the international field and give guidance to those who are willing to work within this limited experimental research work. The standard of dissertations should not be allowed to disregard internationally acknowledged levels of quality. The clearing committee could express positive appraisal of a dissertation in an English foreword signed by the chairman.

The clearing committee itself should seek research connections with neighboring sciences and could thereby find support and control from outside.

It is evident that these guidelines mean an enormous amount of restriction for the individual researcher, but they were written with the conviction that the time of wandering over colorful meadows, regarding flowers here and butterflies there, turning from one direction to another should now be over. A fence for future activities will certainly be helpful.

References

- BERGLER, R. (1975), *Das Eindrucksdifferential. Theorie und Technik*. Bern.
- BLALOCK, H.M. & A.B. BLALOCK, eds. (1968), *Methodology in social research*. New York: McGraw-Hill.
- CAMPBELL, D.T. & J.C. STANLEY (1963), *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- CONNER, C. (1987), Learning styles and environmental knowing. In: *Tagungsbericht und wissenschaftliche Abhandlungen, Deutscher Geographentag*. Wiesbaden: Steiner Verlag, pp. 482-487.
- DUNKIN M.J. & B.J. BIDDLE (1974), *The study of teaching*. New York: Holt, Rinehart and Winston.
- EGAN, D.E. & J.G. GREENO (1973), Acquiring cognitive structure by discovery and rule learning. *Journal of Educational Psychology* 64, pp. 85-97.
- GAGE, N.L. (1978), *The scientific basis of the art of teaching*. Teachers College Press, New York.
- GRAVES, N.J. (1984), *Research and research methods in geographical education*. London: University of London Institute of Education.
- GRIZZLE, J.E., C.F. STARMER & G. KOCH (1969), Analysis of categorical data by linear models. *Biometrics* 25, pp. 489-504.
- HAUBRICH, H. (1984), Research methods in evaluating the use of media in the teaching of geography. In: Graves 1984, pp. 119-142.
- HEILIG, G. (1982), *Schülereinstellungen zum Fach Erdkunde*, Dissertation TU Munich, also as: *Geographiedidaktische Forschungen* 10, 1984. Berlin.
- HEILIG, G. & H. SCHRETTENBRUNNER (1982), *Erfahrungsbericht: Empirische Forschung in der Didaktik der Geographie*. *Geographie und ihre Didaktik* 3, pp. 114-133.
- HOLM, K. (1974), Theorie der Fragenbatterie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 26, Nr. 2, pp. 316-341.
- JOHNSON, D.W. (1981), Student-student-interaction, the neglected variable in education. *Educational Researcher* 10, pp. 5-10.
- KLAUER, K.J. (1973), *Das Experiment in der pädagogischen Forschung. Studien zur Lehrforschung* 3. Düsseldorf.
- KLAUER, K.J. (1985), *Framework for a theory of teaching*. *Teaching and Teacher Education* 1, pp. 5-17. London.
- KUCHLER, M. (1979), *Multivariate Analyseverfahren*, Teubner Studienskripte. Stuttgart: Teubner.
- LEUTNER, D. (1988), *Entdeckenlassendes Lehren*. *Geographiedidaktische Forschungen* 17, pp. 24-43.
- LEUTNER, D. (1991), *Implementation und experimentelle Evaluation von Lernhilfen im computersimulierten Planspiel "Hunger in Nordafrika"*. *Geographiedidaktische Forschungen* 18, pp. 112-136.
- LEUTNER, D. & H. SCHRETTENBRUNNER (1989), *Entdeckendes Lernen in kognitiven Realitätsbereichen*. *Unterrichtswissenschaft*, pp. 327-341.
- NUSSBAUM, A. (1984), *Entdeckendes Lernen: Probleme der Forschung und mögliche Lösungen*. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* XVI (1), pp. 57-76.
- NUSSBAUM, A. & D. LEUTNER (1986), *Die Auswirkung der Schwierigkeit textbegleitender Fragen auf die Lernleistung*. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* XVII (3), pp. 230-244.
- ROLLETT, B. (1969), *Das Design in der empirischen Unterrichtsforschung*. In: L. Roth, *Beiträge zur empirischen Unterrichtsforschung*. Hannover: Schroedel.
- ROLLETT, B. (1978), *Lernpsychologische Untersuchungen als Grundlage geographiedidaktischer Planungen*. In: H. Schrettenbrunner, ed., *Quantitative Didaktik der Geographie, Teil II, Der Erdkundeunterricht* 28, pp. 39-55.

- ROSENSHINE, B. & N.F. Furst (1973), The use of direct observation to study teaching. In: R.M.W. Travers, ed., - Second handbook of research on teaching. Chicago: Rand McNally, pp. 122-183.
- SCHRETTENBRUNNER, H. (1984), Empirical methods for researching media - A challenge for more awareness of methodology, 25th International Geographical Congress Freiburg 2, pp. 489-499.
- SCHRETTENBRUNNER, H. (1986), Research perspectives for the empirical didactics of geography. In: Abstracts and Papers, Commission on Geographical Education IGU: Geographical Education and Society. Sitges: IGU, pp. 46-59.
- SCHRETTENBRUNNER, H. (1988), Memorandum for a working group: Empirical research in the didactics of geography. In: R.Gerber & J.Lidstone, eds., Proceedings of the Skills in Geography Education Symposium, vol 1. Brisbane: IGU, pp. 322-332.
- SCHRETTENBRUNNER, H., ed. (1989 and 1991), Software für den Geographieunterricht. Geographiedidaktische Forschungen vol. 18. Nürnberg (1991): Lüneburg (1989), Hochschulverband für Geographie und ihre Didaktik.
- SCHRETTENBRUNNER, H. (1990), Results of the working group on empirical research. In: Teaching Geography in and about Asian Pacific Countries. Hong Kong: IGU, pp. 309-314.
- ZIMMERMANN, E. (1972), Das Experiment in den Sozialwissenschaften. Stuttgart: Teubner.

2 THE DEVELOPMENT AND EVALUATION OF A GEOGRAPHY COMPUTER SIMULATION

Helmut Schrettenbrunner

2.1 Basic assumptions for the development

In Germany professors and researchers in the didactics of geography are often specialized in the development of teaching material or media. Thus the university pays its contribution to the practical applicability of university contents at the school level and hopefully reduces the time lag between two different types of educational institutions. Admittedly, the professional production of such teaching material or media is in the expert hands of specialized publishing houses, and the researcher is only one figure among others in the editorial group which is responsible for the proper transfer of a) new contents, b) new methods (instruments of geography) and/or c) teaching strategies.

In Germany these three components have undergone fundamental changes since the late sixties due to the emerging subdisciplines of geography (social geography, ecology, applied and urban geography, regional planning, just to mention the most important of them), to altered concepts of education and teaching (such as discovery learning, CAI, etc.), and different technical equipment in school (recorders, projectors, computers). Large efforts were made in the early seventies to implement the new ideas, and the American High School Geography Project (HSGP) played an important role (Geipel 1969) and was subsequently imitated and altered by a German project (RCFP, final report Schrettenbrunner 1981) which even today is remembered by a small group of teachers (Niemz 1989, p. 155, 19% remember RCFP).

It was as late as the seventies that games and simulations became fashionable in geography, propagated again by US examples (e.g. the HSGP: the Game of Farming, the Railroad Game) or by British publications (Walford 1969, in a German translation Walford 1972 with the North Sea Gas Game, the Bus Line Game, etc.), which lead the way to own productions, such as an Airport Location Game (Franz et al. 1977) or an Urban Renewal Game (Schrettenbrunner 1973), just to name the first. And it was implicitly assumed that the new games, media and teaching methods would give strong impulses to geography in school and would reduce teacher monologues and diversify teacher-student dialogues. Certainly, the state curricula were changed as a result of these efforts and permitted a "new geography". The production of new teaching materials kept the publishers busy for a decade. But after 20 years we have, however, to admit that the impact of games or simulations is evidently very low in school life (Niemz 1989, p. 139, 31% of the young teachers interviewed stated that they "occasionally" use them, but only 18% of the older teachers).

There may be much speculation about this low interest in games and simulations

in German geography classes. As the educationalist and specialist for games Einsiedler (1989, p. 298) put it "learning or didactical games have been questioned again and again to ascertain if they can be brought together with intentional teaching, and what function games should have in classroom activities". He recommends play oriented learning methods in certain domains as a reasonable complement to the teaching-learning-processes in school because "they enforce cognitive components in the process of playing and still have motivational and emotional advantages"... "and help to develop the student's personality" (p. 305). Evidently, German geography teachers who have dozens of games at hand do not accept this point of view, and Niemz in his analysis can only guess that despite all innovative efforts the games just were not convincing enough (1989, p. 141). It may be added that cardboard games or role plays need an expert who does not give up after a few initial failures. Little evaluation has been done in this field of geography teaching (see for a rare example Schilke et al. 1984 on a role play which is related with our simulation: agriculture in Africa), so that no one really knows the reason for the low level of acceptance.

Now we can make things even more complicated if we combine simulations with computer games, bearing in mind that at least in 1988 the practical use of computer in geography classes was virtually zero (Niemz 1989, p. 134). It should, of course, be considered as a real challenge for the didactics of geography to help introduce an innovation when evident obstacles (teacher's qualifications, inservice training, learning situations, organizational problems) must be overcome.

We have to take into account that the student's activity is different when he works with a computer simulation. He encounters a very complex dynamic model and must develop strategies to solve a problem, in our case running a farm. He can manipulate some variables and should witness the effects on development variables, thus learning the implicit rules of the game.

He can do so several times (rounds, years) to explore the model situation. He can start again and again without fear of personal consequences (such as a bad grade for geography) and should be able to identify himself with an overall aim (in our case: keep the family from starvation). This kind of learning can be seen in the tradition of "problem solving" and "discovery learning", which may have had more impact in educational research than in everyday geographical classes. To the geography teacher an important question is to what extent a student with little experience in discovery learning will cope with such a computer program. Furthermore, will he be able to learn anything at all, if he is not an expert in the subject content as well? Probably the implicit fear of our geography teacher is now clearer than before and might be verbalized as "a simulation is nothing but a game". The educationalist Klauer (1985) describes it like this: there is an evident lack of teaching functions in most simulations which may prevent solid learning results. Such functions - according to Klauer (1985, p. 10) are "five molar teaching functions: ... motivation, information, information processing, storing and retrieving, transfer of information, monitoring and directing"). A normal computer simulation does not offer these functions but relies on individual exploration and numerous repetitions. The student can, of course, learn that the simulated system changes in the course of time (people in the simulation get older every year/round, events happen at a certain point of time, etc.), or events occur as a result of his activities (plants grow after his selection, agro-technical investments he made help to produce more), or things happen according to a random selection (climatic events occur, social or political

mishaps come and go) or according to a mixture of student's actions and system situations. Altogether, this may describe the very complex model of a simulation, and again it can be questioned if high complexity - which is necessary for a meaningful simulation - is conducive to good learning results. Certainly, our geography teacher will ask in addition how much time can be allotted to such activities, which, indeed, is a very important and limiting factor in the classroom.

Bearing such crucial points in mind we decided in 1985 to concentrate (again, after a first computer program for an experimental school in 1971, see Schrettenbrunner 1973) on the development and evaluation of computer programs, because we wanted a solid basis for an empirically orientated didactics of geography. Our major aims have been:

- to develop computer simulations in close connection both with students and teachers in schools (as later users) and researchers in universities (as geography experts);
- to build up a group of users/sponsors and administer the spatial diffusion of our programs ourselves;
- to optimize our programs according to the reality of the classroom, testing different methodological approaches or versions with the help of our users;
- to analyze students' learning effects and reactions to our programs using theoretical concepts which normally are found outside the didactics of geography, particularly in educational science, psychology or pedagogic.

In short, we consider the didactics of geography to be an applied science, with the aim of experimentally analyzing teaching material with geographical contents/methods (based on theoretical concepts of geography) and optimizing teaching methods or learning strategies (using theoretical concepts of other behavioral sciences).

2.2 "Hunger in the Sahel", a simulation

Admittedly, the choice of our subject was not very original. In fact, it turns out that this theme had been chosen before by a number of software authors. They all came up with ideas, gave inspiration, and above all, urged us to alter presentation and strategies. For example, we did not want to accept the macro-economy approach presenting a large tribe in the arid zones of Africa as most authors do (among the first: Doerner & Reither 1978 with Tanaland and Moro; Baumann 1985 with Tangaland; Koschwitz & Wedekind 1985 with Dori; DIF 1989 with Nara), because their main interest may have been more in economy than in behavioral geography. We decided to adopt the micro-economy approach presenting one family with their possible aspirations and activities (close to the Dutch version "Honger of erosie" by Jutten 1985), because we wanted to integrate the behavioral approach typical for social geography. Furthermore, we assumed that it would be easier for the student to identify with one family, to plan his decisions, and note the consequences on his ten lots of land, rather than to deal with a population of 10,000, a large area which cannot be presented on the screen and with no indications as to why decisions were made.

We also decided that the main menu should look like a vertical cut through the landscape, showing the fields, the plants (as icons), the wells and terraces (also Jutten 1985) and not just statistical lists or diagrams as the other programs do. An

old geographical tradition which goes back to the first half of this century was revived again on the computer screen, the landscape and land use profile.

And a third factor was considered to be essential for a good simulation: We built our model after a real case described in a variety of articles on the Sudan, specifically one tribe in North Darfur. This is in strong contrast to all other simulations on our theme which use an overall economic model with little connection to African reality. The same applies to the partly involved ecological models, too. We had the chance to get the author of these articles (who, furthermore, comes from the Sudan and now is a geography professor in Germany) to summarize his findings on "our tribe" for a joint publication (Ibrahim 1989), to help us with the simplifications in the computer model and to adjust the algorithms.

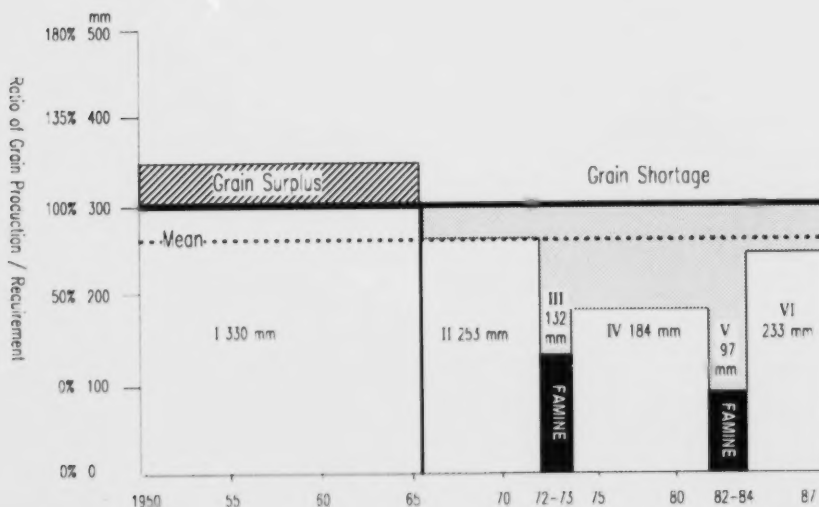


Figure 2.1 Rainfall and drought in the Sahel (from Ibrahim 1988a, p. 135)

To summarize, we built our simulation around solid research on a real situation (regional geography of North Darfur), integrated an important theoretical approach (behavioral geography), adjusted all of the algorithms of the model by expert rating (subject matter validity), and developed a graphic presentation on the screen resembling a traditional landscape profile which shows the student every 'decision' he makes and every effect he causes (after harvest), partly on the lot, with icons or colors, partly as figures and sums.

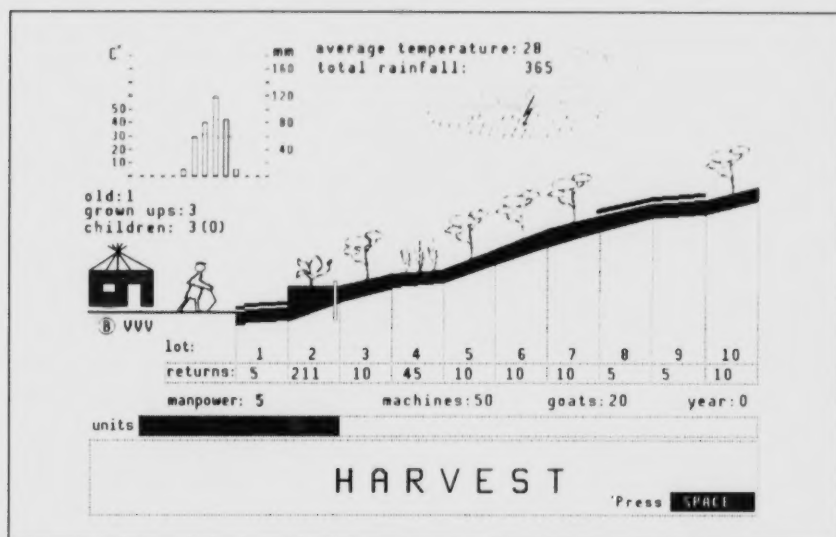


Figure 2.2 Main menu of the "Hunger in the Sahel" simulation (English version)

2.3 Strategies in the simulation

At the beginning the student is told that he is to manage a farm in an arid zone of Africa and is informed that two major risks must be tackled: too little rainfall over the whole year (drought) resulting in a poor harvest, or too much rain in a short period (thunderstorm) resulting in soil erosion and poor land in the long run. He is also told that he can consult information pages on the screen.

He can then make general decisions concerning family planning, the education of children, mechanization and cattle breeding or can plant crops on a slope with ten fields (three grades of steepness), build wells, terraces and grain storages. His decisions are limited by the amount of money and the manpower available (as a function of family size and mechanization); there is no time limit. After all that he can start the calculations ("harvest") of the simulation:

- climatic events (over the whole year), real climate data from Khartoum (which follows a set order of years for our experiment, otherwise there is a random choice);
- harvest results on each lot, depending on the crop, the climate, irrigation, terraces, erosion, etc.;
- erosion damage resulting from rainfall, crops, terraces, erosion of pre-years, steepness of lot, etc.;
- cattle/goats farming;
- family events, births, deaths, marriages, social festivities, nutrition, manpower (with reduction after malnutrition), migration;
- net remains (ready for the next year/round) and a numerical value for the overall result (total value of the farm, comparison to last year).

So the simulation handles 3 models:

- a) an agricultural farm where decisions concerning land use and cattle breeding are made, where harvest incomes can be guessed in advance and real incomes are calculated according to set factors,
- b) an ecological situation with climatic events, soil condition and technical investments, such as terraces, wells, mechanization, and
- c) a family model with natural and social population movements.

For the player there are different strategies feasible which could be differentiated as follows:

Strategy 1

Cattle breeding and soil conservation

The player has decided to specialize in animal breeding and uses his total area as grazing fields, avoiding all of the problems of soil erosion on the slope. He exploits his savannah bushland, especially in drought periods and makes extra income from timber, fireplace wood, plants, mushrooms. Ibrahim (1989, p. 64) describes the real situation in the Sahel: "There are strategies for survival in critical years such as the traditional forms of collecting the fruit of various trees, the milk of rubber trees, fodder for animals, timber and dry grass as construction materials". Cattle breeding requires certain minimum areas for pastures, and in the simulation animals can die of malnutrition as well as of a lack of water. Our player in this game must find out this balance of cattle per lot, cattle per well, and a maximum number of cattle per farm and still he will have a very low income more typical for nomads than farmers, an income which will prevent him from sending his children to high school.

Strategy 2

Cattle breeding plus dry farming

Millet is the traditional crop of the zone. It does not need much water and does not need deep ploughing or technical investments on the one hand, but does not produce high incomes on the other hand. This strategy comes very close to reality. Ibrahim (1989, p. 61) describes this type: "After cattle breeding the growth of millet is the most important thing. Generally speaking, the wife takes care of about 4 hectares of millet, and husband and children occasionally help her. On the average a family produces only 50% of millet needed for their own consumption over the year (which totals 1200 kgs). There is a certain risk not only in drought periods but also when pests, insects, birds or other animals endanger the harvest in wet years".

Strategy 3

Cashcrops and mechanization

If there is enough capital and manpower the growth of cash crops (vegetable, cotton) promises good incomes, running the risk of high ecological damage, if no measures against erosion have been taken. Irrigation is indispensable, which soon, however, leads to an over exploitation of underground water and to an ecological collapse in short period of time. Because our player needs to make many investments he tends to sell all of his animals, with the result that his nutrition causes more costs and that he will have to buy fertilizers. In our experience most students try this strategy first and want to maximize income until they notice that this European approach usually ruins the African system within a few rounds.

Strategy 4

Traditional mixed farming

The traditional way of running a farm is a mixture of cattle breeding, millet growing, with a few wells or water holes, a crop storage and some terraces and a cashcrop (in this order, if money is short) to ensure a sufficient basis of nutrition and avoid ecological damage. Further investments such as higher education for the children can only be made when the farm has survived the first and crucial years and when enough savings have been made. "Only by diversification of his agricultural decisions can the farmer survive" (Ibrahim 1989, p. 61).

The major objective is to make the student understand that in a zone of natural hazards a solid strategy for survival must have the following components:

- the prevention of ecological damage;
- the minimization of risks by diversification of agriculture;
- having an emergency supply;
- keeping down the number of children.

"The traditional economy of quasi subsistence may be considered old fashioned and out-dated by many planners for the Third World but even today it is the only chance for survival in the arid zone of the Sahel when droughts occur year after year" (Ibrahim 1989, p. 65).

2.4 Experimental evaluation of different versions

More or less at the end of the construction phase for our simulation "Hunger in the Sahel" and at the beginning of the informal testing with single students or small groups we encountered questions which seemed important for further research:

- 1) Do students really learn anything that can be classified with cognitive objectives?
- 2) Is the handling of the simulation simple enough for our target group (class 7)?
- 3) Is the tutorial (a summary of the major issues of the simulation) at the beginning helpful to students?
- 4) Are the warnings and explanations given during the simulation useful for the learning process?

At this time (1987) we did not yet know that this was the beginning of a series of experiments which were being conducted by different researchers, in different countries (Germany, Austria, Denmark, Holland) and in various types of schools (Hauptschule, Mittelschule, Gymnasium, University). All those results that could optimize our simulation for its use in school were consequently incorporated for the benefit of our users and sponsors.

The following chapters are summaries of the findings educationalists at the Aachen Institute of Educational Science made as well as presentations of our own research.

2.4.1 The implementation of teaching functions (Leutner 1989)

If we take the criticism to heart that "normal simulations" do not constitute an efficient teaching method because they do not structure learning situations so that the students must compensate - mostly by numerous repetitions - the deficient structure

(Wager & Gagne 1988, p. 46: "... simulations by themselves are seldom an efficient instructional methodology. Most lack important events of instruction like recalling prerequisite skills, providing structured learning guidance, and giving corrective feedback") the constructor of a "better simulation" has to build in a number of interfaces:

- an expert interface of how to steer the simulation, in our simulation only two pages of "rules", a numerical "cost-benefit-list" and short "explanations";
- a tutor interface which should compensate different entrance levels of knowledge, in our simulation eight pages about climate, soil, agricultural, families and regional risks (for the experiment in the form of a linear learning program);
- a diagnosis interface for on-line warnings, corrections or explanations to help the student discover the implicit rules of the model and prevent him from getting stuck or moving in useless loops. Our simulation has over 20 adaptive comments as warnings ("if you dig too many water holes the ground water level may collapse"), corrections ("too many goats on your pastures, do you want to send them to distant pastures? Or take another action?"), and explanations ("Drought! Not enough water for the plants, harvest will be poor without irrigation").

If all three interfaces were fully developed we would be justified in calling our simulation an intelligent tutoring system, but so far we have (with the cooperation of Leutner) only elaborated on the diagnosis interface and are still experimenting with the tutor interface, and to a lesser extent with the expert interface.

Most of the deficient teaching structures Klauer (1985) mentioned (presentation of teaching material, elucidation of implicit information, teaching aids and guidance) could be corrected for the first experiment.

Leutner (1989) wanted to test the effects of "presentation of information" and "elucidation of implicit rules" on learners' results. "With regard to the presentation of information we wanted to find out if the learning task of doing the linear program (on fundamental concepts, rules and principles of the simulation) before the game is helpful for the subsequent simulation (Ausubel 1960; Tennyson & Christensen 1988; Anderson 1987). Furthermore we wanted to find out if the elucidation of implicit rules and control of learner's actions had been successfully built in with our adaptive warnings and comments which during the simulation turn up depending on the situative state of the simulation as well as on the moves that the player has made so far. All this should provide positive learning effects (literature concerning intelligent tutoring systems: Kunz & Schott 1987; Mandl & Fischer 1985; Mandl & Lesgold 1988)" (Leutner 1989, p. 336).

We developed four different versions of the simulation, "with vs. without" a tutorial program at the beginning, and "with vs. without" adaptive learning aids (2 x 2 factorial design), our hypotheses being that both learning aids should have positive effects on learning results. A cognitive criterion-referenced test (pretest, posttest) was developed, and an intelligence sub-test (Horn 1962) as well as a personality test (Schwarzer 1986) were employed. 64 students in grade 7 participated in the experiment (16 per experimental case, randomly attributed).

The procedure was as follows:

- 1) Cognitive pretest to measure learner's entrance level, IQ-test, personality test.
- 2) Experimental treatment (after 1 week).
 - a) Exploration of the 4 versions (45 mins.) with new starts or interruptions at will.

- b) Action-test (max 30 mins.). Here the students were told to give their best ("How long can your farm survive?") and no repetitions were allowed.
- 3) Cognitive posttest (after 1 week).

The findings can be summarized as follows (Leutner & Schrettenbrunner 1989, pp. 337-339):

- The simulation provides a good learning situation. Leaving the four versions outside the calculation, a pretest mean of 15.6 and a posttest mean of 27.5 prove that the students learned something (which can be described as normal geographical knowledge) in a reasonable time;
- Posttest results (cognitive domain) can be differentiated from action-test results (to guide the simulation), the correlation of $r = 0.02$ shows that the two variables are independent;
- Learner's prerequisites were controlled by step by step multiple linear regressions;
- To measure the effects of the four versions analyses of covariance were calculated and the following results obtained. There is a significant main effect of the adaptive aids on the cognitive posttest ($F(1,54) = 5.43, p = 0.024$) as well as a significant interaction of tutorial and adaptive aids ($F(1,54) = 9.23, p = 0.004$) which means: The version with adaptive aids leads to better posttest results than the version without adaptive aids, and the maximum of posttest results is provided by the version "with adaptive aids but without tutorial";
- There is no significant effect concerning the action test and the experimental version, only an interaction between adaptive aids and tutorial on a weak level of significance ($F(1,54) = 3.98, p = 0.051$).

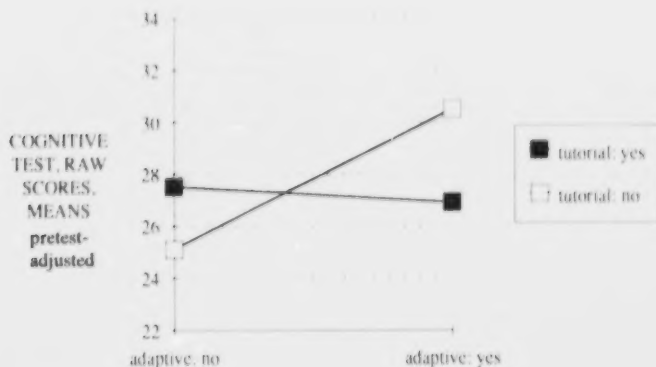


Figure 2.3 Cognitive test: experimental conditions (adaptive vs. tutorial) (from Leutner 1989, p. 87)

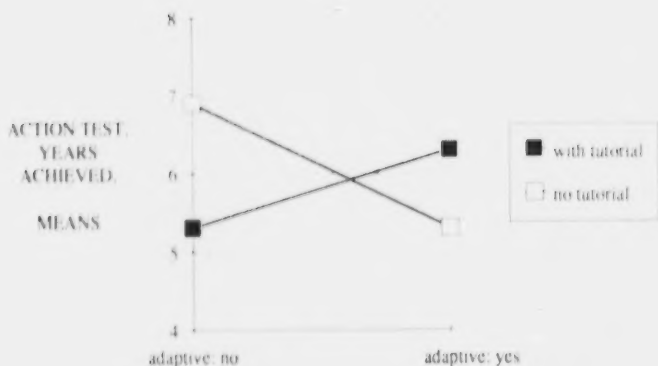


Figure 2.4 Action test: experimental conditions (adaptive vs. tutorial)
(from Leutner 1989, p. 87)

Leutner & Schrettenbrunner (1989, p. 339) give the following summary and discussion: "The findings of our evaluation prove that the computer simulation *Hunger in the Sahel* is an effective means of teaching. It has been a surprise that the two aspects of learning effects (cognitive test vs. action test) are independent, which is interesting from a methodological point of view. Those students who had to explore the simulation under a strict discovery method (no adaptive aids, no tutorial) got the best results in the action test (i.e., they really could guide the simulation) but got significantly lower results in the cognitive posttest. In contrast, those students who could dedicate their whole time to the adaptive version and did not need to do the tutorial gained the most points in the cognitive test, but the least in the action test. Once again: one group just learned to play the game successfully, while the other group learned more geography, which for school purposes certainly is more desirable. Our recommendation for the use of computer simulations in school is therefore, to use only those with adaptive aids built in, unless there is a tutor per group who could help".

Accordingly, we distributed the adaptive version to our users but left the tutorial as an option at the beginning now sequenced in six chapters which can be chosen individually (climate, soil, agriculture, families, cattle breeding, general risks) and consist of 1-2 pages. We made the research results available to all of the users (it is, in fact, now an article included in the manual) and gave advice on how to build the simulation into a normal sequence of teaching units. The teacher, in fact, should present the contents of the tutorial himself as a kind of preparation. He should use the atlas or visual media which the computer cannot offer, and then let the students explore the simulation for one full lesson/hour to get the maximum results, always assuming that time is the limiting factor in school reality. After that he should continue with a discussion to find out to what extent the major concepts (drought, erosion, irrigation), main problems (scarcity of means, decisions against better knowledge), or general geographic insight (land use, concepts of behavioral geography) were

conveyed. We also recommend strongly to use other geography media (slides, films, books, maps) and to regard the computer as just one medium among others. If the didactics of geography is considered an applied science it should be able to give advice which can at the same time be regarded as a strategy by which to disseminate an innovation.

The feedback we get from our group of sponsors/users is another instrument to continue the dialogue with our teachers, too, while workshops which are financed by the software income are another means of getting teachers and researchers together. In our opinion, we need to do more than just develop software. We also need to organize the dissemination and build up a new kind of interaction between university and school.

2.4.2 Checking the function of the adaptive aids (Leutner 1991)

One result of the first evaluation caused us to check the function of the adaptive aids. These aids help to get good results in the cognitive posttest (if the student has enough time to explore the model), but they seem to hinder the performance in the action test (steering the simulation). Certainly, the correlation coefficient of $r = 0.02$ between cognitive and action test was a surprise, too.

As this was a question in which the "tutorial" was not involved a simple design was planned and 2 versions were tested:

- a) exploration without adaptive aids and
- b) exploration with adaptive aids.

Thirty-eight students of the Aachen University of Technology took part in the 2-group design with a similar procedure as before, but all experimental activities were conducted in one day.

Leutner (1991, pp. 158-159) describes the results as follows: The adaptive version produces better results in the cognitive posttest (raw point means are 32.2 respectively 30.0 which is significant on a 4.2% level). The differences in the action-test are lower, as was tendentiously expected, but not significant (raw point means are 4.8, respectively 5.8, $p = 0.068$). "The aim of this experiment was to replicate findings of the first evaluation. The replication can be considered successful: Adaptive aids of this kind in complex dynamic computers simulations can have the effect of advancing cognitive verbal and general knowledge but hindering action and simulation specific competence. This effect can be noted with 7th grade students as well as with young adults. A question which at this point is still unclear is which of the components of the adaptive version is responsible for these effects in the end" (Leutner 1991, p. 159).

2.4.3 Optimizing the non adaptive aids (Leutner 1991)

In the simulation "Hunger in the Sahel" there are different learning aids such as:

- adaptive comments
- a tutorial (non adaptive)
- background information (non adaptive cost-benefit pages).

These pages with background information can be called up by the student at any time during the simulation to get information about the cost of an action and the benefit from an action, but in the adaptive version they also offer explanations (reasons for, effects of measures). Much depends on the students eagerness to

consult these pages and accept this teaching aid and as quite a lot of figures are involved how often he needs them may also depend on his memory. It was obvious that the amount of information as well as the kind of presentation (screen vs. paper) might also influence learners' effects. Four different versions were subsequently developed by Leutner, which can be described as follows:

- Strict discovery type, equal to the former "without tutorial, without adaptive helps", only with the distinction that the cost-benefit pages were removed from the simulation and placed as a photocopy next to the keyboards. This sheet no. 1 was used in all four versions.
- Discovery type as above but with a second page of the cost-benefit pages giving background information and explanations, such as "Bushland preserves the slopes from erosion, also useful as pasture. Reforestation regenerates soil". There was more of this kind of information compared to the first experiment.
- Adaptive type as in the former experiment. The adaptive commentaries were enlarged to a total of now 34. Sheet no.1 was added.
- Adaptive type with two pages of cost-benefit information.

INFORMATION

Values change during
the simulation!

decision	manpower required	costs	returns
bush	10	10	10
pasture	0	0	5
millet	10	20	150
cotton	20	30	120
vegetables	25	40	150
terrace	10	50	(increase)
well	10	50	(increase)
1 goat (max. 2 goats/pasture)		10	10
machines		equal	(no sale)
school		100	(later)
family planning		- - -	- - -

nutrition: old 25 / grown ups 50 / children 40

manpower: old 5 / grown ups 10 / children 5

Next: press SPACE

INFORMATION

Values change during
the simulation!

decision	notes
bush	- good against erosion, pasture for goats. reforestation good for soil regeneration.
pasture	- graze land for goats.
millet	- simple, undemanding cereal.
cotton	- needs much water, gives high turnout.
vegetables	- highest turnout, needs much manpower.
terrace	- against erosion.
well	- better returns, water in times of drought, without well goats will die in drought.
machines	- give more manpower, 10 units machines = 1 unit manpower.
school	- for jobs outside agriculture, additional income for family (later)
family planning	- few children = less manpower, less foodstuff, but also less old age support for the parents.
millet- granary	- costs but helps in time of drought.

Next: press SPACE

Figure 2.5 Cost-benefit pages (English version)

The procedure was the same as described earlier. Eighty students in grades 7 and 8 took part in the experiment (2 x 2 factorial design, 20 per experimental case, randomly attributed). After the posttest a retention test was administered 2-3 weeks later. Dependent variables were action test, posttest (2 parts) and retention test (2 parts). Results of the multivariate analysis of variance are shown in the following table (Leutner 1991, p. 162).

BEST COPY AVAILABLE

Table 2.1 Multivariate analysis of variance

Source		Lambda	F	df	p
Covariates					
IQ-subtest (verbal)	(V)	0.87	1.96	5,64	0.096
inductive thinking	(I)	0.87	2.00	5,64	0.094
anxiousness	(A)	0.79	3.50	5,64	0.008
self assuredness	(S)	0.95	< 1		
Experimental treatment					
adaptive comments	(C)	0.93	1.03	5,64	0.410
background info.	(B)	0.80	3.20	5,64	0.012
interaction	(CxB)	0.95	< 1		
ATI-effects					
(IxC)		0.96	< 1		
(IxB)		0.90	1.46	5,64	0.216
(IxCxB)		0.86	2.08	5,64	0.079
(AxB)		0.64	7.24	5,64	0.000

Leutner (1991, p. 62) summarizes the main effects which can be traced back to the experimental manipulation in the design:

- 1) There is a strong and highly significant main effect of the cost-benefit-list ($p = 0.012$) which prove the effectiveness of such background information. They are more important than the adaptive comments.
- 2) There is a similarly strong ATI-interaction between "anxiousness" and the "cost-benefit-list" which means that students who are less anxious profit more from this background information.

Univariate analyses of the means render the following results. The importance of the photocopies next to the keyboard with important background information can be shown on a significant level only ($p = 0.001$) in the retention test, see figure 2.6.

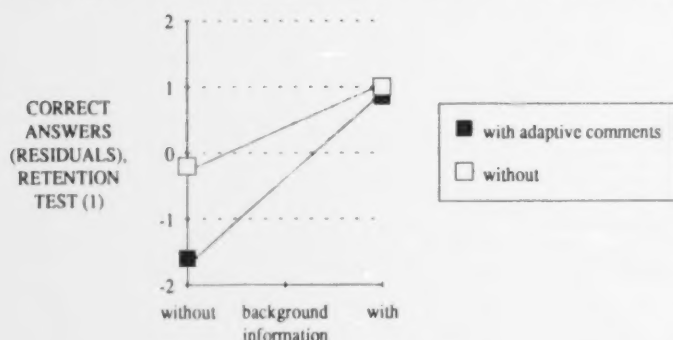


Figure 2.6 Correct answers in the retention test (Leutner 1991, p. 165)

This is not evident in the posttest, although there is the same tendency. Leutner (1991, p. 166) discusses his findings placing emphasis on the importance of the permanently present background information. In a complex computer simulation this may be a satisfactory learning aid because the student can peruse the whole cost-benefit-list again and again and may get a kind of mastery learning effect. He may also have come to a deep level of learning in the sense of the "levels of processing approach" Craik & Lockhart (1972) described. Students may have noticed that the background information gives them the key to the simulation with all the rules and explanations, and they have learned their lesson better. Both concepts (mastery learning and levels of processing) lead to the expectation that such background information will be remembered with a long term effect. And indeed, the importance of the cost-benefit list cannot be proved in the posttest, only in the retention test.

Compared to the importance of background information the adaptive component (comments) is secondary but it helps in the same direction. No significant effects could be found for background information on the action test. It may be recalled that in the first experiment a negative effect of the adaptive version on the action test was noted.

These findings gave us occasion again to alter the new edition of our manual for the simulation (1991 edition). We included four pages with all the rules and explanations (cost-benefit list, differentiated according to the different levels of difficulty) and advised the teacher to photocopy them for each computer group. The pages on the screen have still remained in the program, especially for the more difficult versions when values are changed in the course of the game according to the state of the model (adaptive changes).

As to the lower importance of the adaptive comments it could be speculated that one would have to play a long time to experience all of them and the students who keep starting the simulation again after just a few rounds (in the test version the mean number of rounds is about six) will never see all of the comments. In contrast, the summary of rules and explanations on the cost-benefit list is always at hand, even in details that will not necessarily be needed but can be learned incidentally. Also, we should take into consideration that the visual gestalt of "one page on paper" may lead to better retention effects than that of "two pages on the screen".

2.4.4 Controlling the effects in different school types

The simulation "Hunger in the Sahel" was not designed for a specific type of school and in fact was tested in the lower type (H like Hauptschule), in the upper type (G like Gymnasium) and in others as well. It must be remembered that in Germany there is one common elementary school (grades 1-4) with a subsequent channelling of pupils according to their grades in grade 4 in either the lower or the upper type.

The Gymnasium is the school for

- the brighter children;
- for those children who had more supplementary training by their mothers in the elementary school;
- for those children whose parents consider higher education to be essential;
- for those whose parents would like to send their children to college.

There are highly significant differences between pupils and their learning results when we compare grade 6 students of the different types of school. Research findings which might be interesting in our context point out such differences in favor of the Gymnasium:

- a) a higher level of overall and specific knowledge;
- b) a higher speed in acquiring new knowledge.

Furthermore typical aptitude-treatment interactions can be expected if we compare students from the two types of school in so far as elaborate teaching methods, such as our adaptive version may not be so effective with the better group and may sometimes even hinder their progress (Snow 1980; Snow & Lohman 1984; Leutner & Kretschmar 1988). Klauer 1985, pp. 14-15, puts it like this: "... students high in aptitude learn more efficiently under low-structure than under high-structure settings. If, however, the teaching methods do not require different learning times, the prediction for high-aptitude students would be that the methods will not differ as far as students' learning is concerned". The adaptive version may render better learning results in the cognitive domain and in the action test with students from the lower school type, while Gymnasium students would not need so much guidance and help or might even be hindered in their theoretically higher zest for an explorative situation.

In our experiment we employed our adaptive version of the program "Hunger in the Sahel" and tested it in grade 6 of the lower and upper school type with 60 students in each type. In the first session we administered the cognitive pretest (with some alterations in the phrasing of the items) and IQ-subtest (verbal) and personality test and let the students work for 50 minutes with the program. In the second week we administered the treatment, the posttest (1) and the action test (How long can you keep your family alive?), and in the third week there followed a posttest (2) as retention test and a questionnaire on interest in the program.

Our findings can be described as follows:

- a) There is a marked difference in the entrance level of our cognitive pretest between the school types. It must be added, that the subject matter (the Sahel, erosion) could not have been treated in normal geography classes (it is in fact part of the curriculum in grade 7 and 8). If we keep this in mind, then the difference in pretest results (H: 12,1 and G: 16,4 raw points) means that students in the upper school type may have accumulated more geographically relevant knowledge by themselves than the other group.
- b) There is a marked difference in achievement gradient (from pretest to posttest 1, not so much from posttest 1 to posttest 2 (retention test), which means that the upper group learns more (5.8 points compared to 4.0 points) than the lower group and thus arrives at a much higher level of knowledge and even widens the cognitive distance to the other group.

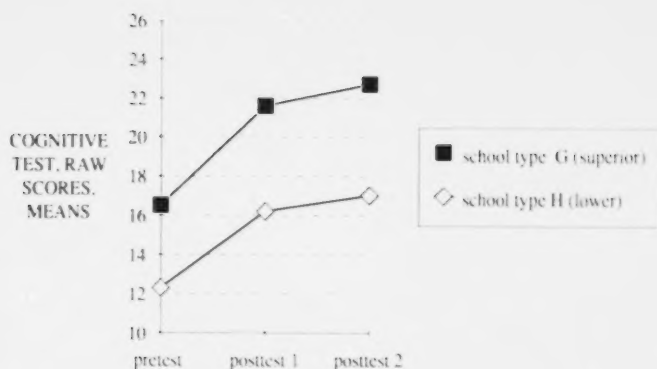


Figure 2.7 Cognitive tests, raw scores, means per school type (experiment 1991)

Table 2.2 Repeated measurement analysis of variance: cognitive tests per school type, level, tutorial, within-subjects effects (see also table 2.4 for between-subjects effects)

	Sum of squares	df	Mean square	F	p
Main effect					
- time (test)	2026.81	2	1013.40	157.91	0.000
2 way interactions					
- school by time	38.12	2	19.06	2.97	0.047
- level by time	41.93	4	10.48	1.63	0.170
- tutorial by time	1.17	2	0.59	< 1	n.s.
3 way interactions					
- school/level/time	16.02	4	4.00	< 1	n.s.
- school/tutorial/time	2.77	2	1.39	< 1	n.s.
- level/tutorial/time	15.63	4	3.91	< 1	n.s.
- school/level/tut./time	11.36	4	2.84	< 1	n.s.
Within cells	1386.20	216	6.42		

It may be astonishing that the results in posttest 2 are even higher than in posttest 1. Two reasons may be given.

1. Posttest 2 is not precisely just a retention test because the students evidently had another teaching experience when they did the action test and played the simulation for another 10-20 minutes. This additional training had a positive learning effect.
2. As most of our students were volunteers we had to motivate them with presents to come 3 times. We explained in the beginning that there will be presents, then we showed the presents in the second session and before the third we repeated that the presents would be given according to test results. This extra motivation helped a lot to avoid drop outs and - in comparison to the first test - lifted learning results.

With reference to the action test and a differentiation according to treatment (30 students per school type did the version with tutorial, 30 without; a 2 x 2 factorial design) the findings are such that there is no difference when the version without tutorial is used, but there is a main effect and a two-way interaction, when the version with the tutorial is applied. Students in the lower school type are less successful in the action test if they had the tutorial before.

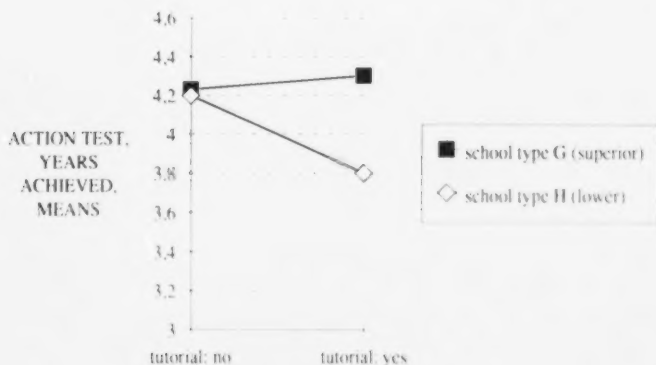


Figure 2.8 Action test, years achieved, means per school type, level, tutorial (experiment 1991)

A discussion of these results repeats well known distinctions between major school types due to the early channelling of students. It could be shown that students of the upper school types have a higher level of knowledge which they must have acquired outside school and geography lessons, that they learn more in the simulation and thus even widen the gap of knowledge between the school types. In both cases it can be confirmed that the simulation provides a good learning situation even for a young age group for whom the simulation is not intended. It could not be proved that students of the upper type have less positive effects when they use the more elaborate version (adaptive plus tutorial vs. adaptive without tutorial).

Table 2.3 Analysis of variance: action test by treatment

	Sum of squares	df	Mean square	F	p
Main effects					
- type of school	2.42	1	2.41	5.17	0.03
- level of difficulty	0.80	2	0.40	< 1	n.s.
- tutorial	1.01	1	1.01	2.17	0.14
2 way interactions					
- school by level	1.27	2	0.63	1.37	0.26
- school by tutorial	1.88	1	1.88	4.03	0.04
- level by tutorial	0.87	2	0.43	< 1	n.s.
3 way interactions					
- school/level/tut.	0.60	2	0.30	< 1	n.s.
- residual	50.30	108	0.47		

If students of the lower school type use the adaptive version with the tutorial (altogether 50 mins.) they are less successful in the action test because they might have had too little practical experience in steering the simulation as the tutorial takes about 15-20 minutes out of the total of 50 minutes. Therefore we would recommend

- to leave the tutorial out in schools where there has been a selection in lower and upper levels and give the students more time to explore the simulation;
- or keep the tutorial and give more time for exploration (another experiment should check the effects).

Snow & Lohman (1984, pp. 372-373) describe the more general instructional implications: "A set of alternative instructional treatments should be designed to vary in 'incompleteness'. For lower general mental ability learners, the instructional treatment should be made explicit, direct, and structured in detail so as to provide the procedural knowledge as well as the conceptual knowledge such learners may not be able to provide for themselves; it should provide guided assembly and control during learning. For higher general mental ability learners, instructional treatment should be left relatively incomplete and unstructured so as to allow idiosyncratic exercise of procedural and conceptual knowledge; assembly and control functions should be unguided".

2.4.5 Installing different levels of difficulty

It may be true that "learning from simulations is mostly a trial and error or discovery learning process" (Wager & Gagne 1988, p. 46) and that simulations do not offer sufficient teaching structures (Klauer 1985, p. 10), but this on the other hand may be the reason why in our own experience simulations offer learning situations for different age groups on a non-defined level of difficulty. We have noticed that children of 10 years of age, university students, school and university teachers can

play the simulation "Hunger in the Sahel" with equal interest, finding an individual level of difficulty in this exploratory situation.

For school reasons and if we want to fit our simulation into the curriculum of geography (though varying in Germany from state to state) we need to produce a version for grades 7 and 8 when African, physical geography or ecological themes are standard contents, and for grades 12 and 13 when Third World problems are also included. Since 1989 we therefore offer the program in four different versions.

- Version A (as has been used in experiment 4.1 - 4.3), with cows as animals, with millet and maize/corn as plants, and financial means of 300 (imaginary \$);
- Version B with goats as animals, with millet, cotton and vegetables as plants, an option of birth-control and emergency supply, financial means of 500 which is higher than before because higher expenditures than before must be expected (clan feasts on births, deaths, marriages, etc.). Versions A and B have a set order of climatic events;
- Version C is like B but with a random selection of weather and certain risks of unlucky combinations of destructive events (drought, thunderstorm) and financial means of 400;
- Version D is like C but with a random selection of social and agricultural events (accidents, mishaps, taxes but also refunds, etc.) and financial means of 300. Version B-D have changing values of costs and incomes according to a market model.

In our opinion a higher degree of complexity is implemented

- by a wider range of activities owing to the expanded model;
- by a random attribution of climatic events and possible catastrophes;
- by social or agricultural events with consequences (cost, loss, gains) and so far not yet included aspects of social and political life (with the exclusion of war);
- by a diminishing degree of financial freedom leading to unrepairable consequences when wrong decisions are made.

We have information which indicates that version A has been played to 220 rounds and we know that the record with version D is only 8 rounds. It certainly is our full intention to have a very difficult version and not let the students believe that he could master it. "If German school leavers believe that they could master the problems in the Sahel after one lesson on the computer the whole point of this simulation would be totally missed" (Ibrahim quoted by Schrettenbrunner 1991). We have accepted this opinion and therefore offer a simple model for basic concepts (version A) where the player can master the model and play indefinitely long, in fact the group who played 220 "years" communicated that they found out our (secret) order of climatic events, our (secret) algorithm for the destruction of savings above a (secret) amount, etc. But we also offer the very complex model (version D) where the player has no chance and will always be beaten by the system. It should, however, also be remembered that the student will not know that there are such fundamental differences in the versions, and that in a normal computer session he will accomplish no more than 5-10 rounds, anyway. A practically hopeless game (version D) can still render a good learning occasion, whereas from a player's point of view it may be without value.

It was obvious that an experiment was needed to test the effectiveness of our decisions regarding the levels of difficulty (experiment 1991).

We expected a simple version to render the best learning results. This would fit in a long tradition of research on programmed instruction (Skinner 1958) or theoretical concepts of psychology of cognition (Eysenck 1977) or approaches to avoid errors during learning (Brown & Vanlehn 1980). Nussbaum & Leutner 1986 analyzed the relationship between the difficulty of questions in a situation of discovery learning and learning effects and came to the conclusion that students learn best when the questions are very simple. "If we apply this to learning with computer simulations we have to ensure very carefully that students are not exposed to exaggerated levels of difficulty" (Leutner 1990, p. 47). If we give the most difficult version, D, to pupils in grade 6 we should expect therefore lower learning results.

We also expect ATI-effects as in experiment 4.4. Students with lower entrance conditions (pretest, school type) should profit more from the simple version and learn less with a more complex simulation whereas students in the upper school type could cope with either version. This kind of interaction would be usual when discovery learning is employed and could be found in various experiments (Egan & Greeno 1973, p. 92, to quote one of the early experiments: "One goal of studying aptitude and instructional variables is to be able to assign subjects of varying ability to optimal instructional methods. The present results suggest that subjects lacking in skills necessary to solve problems may learn more efficiently when instructed by techniques requiring interpretation and application of a rule").

Furthermore, we expected that students would not detect the "theoretical level of difficulty" of their version because the adaptive comments make you feel like you are being taken by the hand and not being left alone (Schrettenbrunner 1976, p. 50, as in the case of a multi-branched geography computer program). We also expected that students would not show different levels of interest in the program.

Our method was to construct three different levels of the simulation:

- a simple one (= version A), cows as animals,
- a difficult one like version D, but without random selection of weather, and without random selection of social and political events, but a set order instead, and goats as animals,
- and a combination of simple (half of the session) and difficult.

The action test had a medium level (like version B). Half of the students were from the lower type of school (60) and half were from the upper type (60), all grade 6 (2 x 3 factorial design with 20 students per cell).

Our findings can be summarized as follows: The level of difficulty is of no importance for cognitive posttest results, but there is a general effect according to school types. Upper school students learn more than those from the lower type of school, independent of the levels of difficulty. And there is interaction between the level of difficulty and the school type. Students in the lower type learn less when difficult versions are used, whereas students of the upper school type learn more when the level of difficulty is higher.

With regard to the action test, no effects correlated to the different levels of difficulty could be found ($F < 1$).

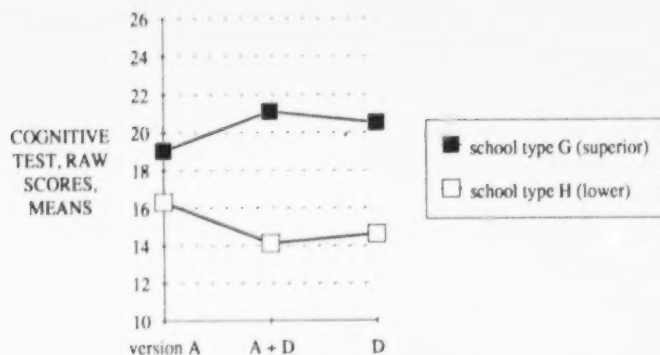


Figure 2.9 Cognitive test, raw scores, ATI-effects between school type and level of difficulty

Table 2.4 Repeated measurement analysis of variance, cognitive test, school types and difficulty (experiment 1991), between-subjects effects (see also table 2.2 for within-subjects effects)

	Sum of squares	df	Mean square	F	p
Main effects					
- type of school	2356.22	1	2356.22	61.57	0.000
- level of difficulty	9.17	2	4.59	< 1	n.s.
- tutorial	3.40	1	3.40	< 1	n.s.
Interactions					
- school by level	408.82	2	204.41	5.34	0.006
- school by tutorial	0.14	1	0.14	< 1	n.s.
- level by tutorial	27.37	2	13.69	< 1	n.s.
- school/level/tut.	133.11	2	66.55	1.74	0.181
Within cells	4133.30	108	38.27		

Students cannot detect or describe the level of difficulty they had worked with. This is a replication of findings we made in an earlier experiment in 1989 with students of grade 8 (in Austria). Figure 2.10 shows surprisingly that in 1991 students who were 1.2 years younger than those from 1989 experienced the versions surprisingly as somewhat easier (the scale is 1 = very easy and 5 = very difficult), but no distinction between our three levels could be found (1989: $F < 1$; 1991: $F < 1$).

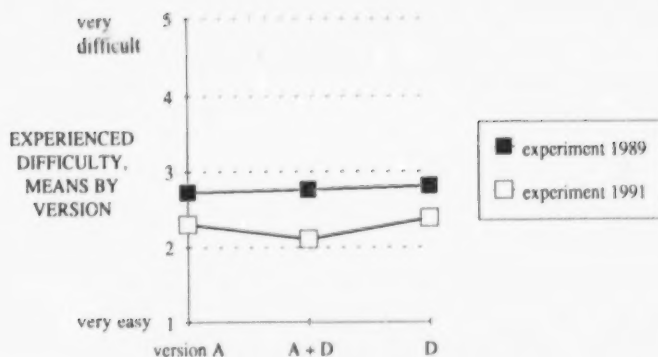


Figure 2.10 Experienced difficulty in experiment 1991 (grade 6) and 1989 (grade 8)

At the end of the three sessions the students answered a questionnaire concerning their interest in the simulation, their experience with computers and other questions. We extracted two factors and called the first "interest in the simulation" and checked factor scores against the level of difficulty. There is, however, no difference. Students are not influenced in their "interest" by the degree of difficulty.

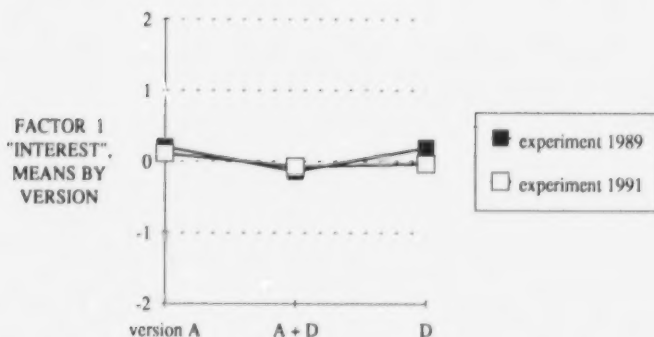


Figure 2.11 Factor 1 "Interest" by difficulty

Table 2.5 Factor analysis, factor matrix, principal components analysis (PC), unrotated

No	Content	Factor1	Factor 2	Factor 3	Factor 4	Factor 5
IT 1	boring-interesting	0.71	-0.18	0.12	-0.27	-0.17
IT 2	not-very reasonable	0.81	-0.02	-0.22	0.07	0.17
IT 3	illogical-logical	0.49	0.13	-0.09	-0.23	0.41
IT 4	embarrassing-clear	0.71	-0.04	0.02	-0.10	0.13
IT 5	dull-motivating	0.23	-0.23	0.25	-0.64	-0.21
IT 6	simple-difficult	0.03	0.32	-0.64	-0.05	0.27
IT 7	beyond-close reality	0.45	-0.13	-0.11	0.70	-0.02
IT 8	nerve killing-encouraging	0.21	-0.28	0.58	0.18	0.43
IT 9	no-lots of comp. games	0.33	0.82	0.18	-0.13	-0.18
IT 10	never-lots of exp. comp.	0.16	0.87	0.19	0.09	-0.09
IT 11	never-often small groups	-0.04	0.22	0.42	0.03	0.53
IT 12	not-love another simulation	0.41	0.02	0.31	0.40	-0.41
IT 13	not-useful for geography	0.79	-0.17	-0.23	-0.02	-0.11

Factor	Eigenvalue	pct.of var.	cum.pct.
1	3.15	24.2	24.2
2	1.80	13.8	38.0
3	1.28	9.9	47.9
4	1.26	9.7	57.6
5	1.05	8.1	65.6

The students have rated the simulation as very interesting (figure 2.12, IT 1), very reasonable (IT 2) and useful for geography (IT 13) and would love to do another simulation tomorrow (IT 12). Learner's motivation is generally speaking very high and there may be different reasons for this (Leutner 1990, p. 43-44; Rheinberg 1985), such as the chance to achieve something, the feeling of immediate success and making rapid progress. The overall positive motivation is fairly constant and could already be found in 1989 with the older group of 8 graders.

A concluding discussion of these results must take into consideration that different levels of difficulty in a simulation will not easily be detected by students. "A learning exercise can be called difficult if there is a great discrepancy between the requirements of the exercise and learner's knowledge. One learning aid can be the adaptive selection of the exercise taking the level of learner's knowledge into account. Another method is to give on-line-help and gradually adjust the learner's knowledge" (Leutner 1990, p. 47). In our case of a simulation with adaptive comments and other learning aids a student may feel he is well tutored and cannot distinguish the theoretical level difficulty.

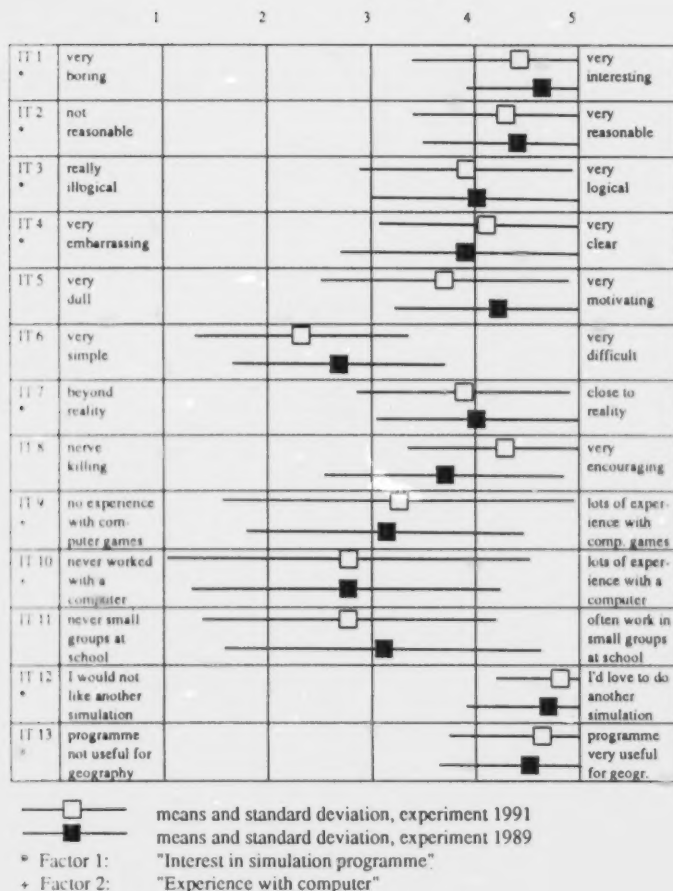


Figure 2.12 Students' attitudes towards the simulation (experiments 1989, 1991)

Learner's results are independent of these levels of difficulty, but advice can be given concerning the different school types again. Students in schools where a downward selection has occurred profit more from the simple version because the evident and simplified model renders an adequate learning situation. They should have the full time of a session to have enough training experience and should not be bothered with a tutorial, or they should have much more time for the exploration of the simulation.

On the other hand, students in upper school types or with better entrance conditions should have the simple version (A) first and then continue with a more complex version (B or C) to let them gradually experience the expansion of the model and notice how the situation becomes more complicated and the effort to survive more strenuous. They might, however, also be confronted with the most complex version (D) right from the beginning without any fear of lower learning results and less positive attitudes towards the simulation.

Our theoretically postulated levels of difficulty have turned out (for upper school types) to be more or less variations (with different animals, plants, options) which can be offered as early as grade 6. As we have seen, the simple version (A) also produces good learning results with students of about 21 years of age (experiment 4.2) with the exception of some ceiling effects when objectives had evidently been internalized before the treatment started. It can be assumed that the simulation "Hunger in the Sahel" is a rather universal medium. Much will depend on the way the teacher integrates the simulation into normal geography lessons and organizes the activities and discussion after the game, thus exemplifying the desired level of difficulty. In our manual we have made suggestions on how to use the simulation in different grades, laying the responsibility in the teacher's hands to decide when he wants to go so far as even to include a discussion on the differences between our four versions, or to elaborate a list of (included or non-included) variables of the model, or to find out our algorithms for the calculation of harvest results, or even to find our secret maximum/minimum thresholds for social events, etc. There is a wide variety of learning events ranging from learning a geographical term to reinventing the construction rules. Users of different entrance conditions and different age groups tend to find their individual level.

Generally speaking, simulations with adaptive aids (e.g., comments on user's decisions) and background information (e.g. cost-benefit pages) offer good learning situations for very different age groups, and theoretically defined levels of difficulty are of minor importance. This is not in concordance with all of the findings of the cognitive psychology or motivational psychology, but it may be true only with this type of simulation. At present it can be hypothesized that these effects turn up only when a simulation has been elaborated in the direction of an intelligent tutoring system (with a minimum of an expert, tutor and diagnosis module). It is the explicit aim of Intelligent Tutoring System (ITS) to implement a universal learner's environment for varying entrance conditions, varying learner's aspirations and varying levels of objectives.

References

- ANDERSON, J.R. (1987), Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review* 94, pp. 192-210.
- AUSUBEL, D.P. (1960), Learning by discovery: Rational and mystique. *The Bulletin of the National Association of Secondary School Principals* 45, pp. 18-58.
- BAUMANN, R. (1985), Tangaland. Lernspiel zur fachübergreifenden Umwelterziehung. In: E. Friedrich & G. Otto, eds., *Bildschirm, Faszination oder Information*. Seelze: Friedrich, pp. 122-127.

- BROWN, J.S. & K. VANLEHN (1980), Repair Theory: A generative theory of bugs in procedural skill. *Cognitive Science* 4, pp. 379-426.
- CRAIK, F.I.M. & R.S. LOCKHART (1972), Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, pp. 671-684.
- DIF (Deutsches Institut für Fernstudien), S. STASCHEN & J. WEDEKIND (1989), *Nara - Leben in der Sahelzone*. Tübingen.
- DOERNER, D. & F. REITHER (1978), Über das Problemlösen in sehr komplexen Realitätsbereichen. *Zeitschrift für Experimentelle und Angewandte Psychologie* 25, pp. 527-551.
- EGAN, D.E. & J.G. GREENO (1973), Acquiring cognitive structure by discovery and rule learning. *Journal of Educational Psychology* 64, pp. 85-97.
- EINSIEDLER, W. (1989), Zum Verhältnis von Lernen im Spiel und intentionalen Lehr-Lern-Prozessen. *Unterrichtsforschung* 4, pp. 291-308.
- EYSENCK, M. (1977), *Human Memory, Theory, Research and Individual Difference*. New York: Pergamon.
- FRANZ, S., G. HACKER & B. KREIBICH (1977), *Im Flughafenstreit dreht sich der Wind*. Braunschweig: RCFP-Westermann.
- HORN, W. (1962), *Leistungsprüfsystem LPS*. Göttingen: Hogrefe.
- IBRAHIM, F. (1988a), Causes of the famine among the rural population of the Sahelian zone of the Sudan. *Geojournal* 17/ 1, pp. 133-141.
- IBRAHIM, F. (1988b), Water harvesting and water spreading für Flutbau in der Sahelzone des Sudan. *Die Erde* 199, pp. 179-184.
- IBRAHIM, F. (1989), Überlebensstrategien der Sahel-Bauern. *Geographiedidaktische Forschungen* 18, pp. 55-66.
- JUTTEN, E. (1985), *Hunger of erosie*. Den Bosch: Malmberg.
- KLAUER, K.J. (1985), Framework for a theory of teaching. *Teacher and Teacher Education* 1, pp. 5-17.
- KOSCHWITZ, H. & J. WEDEKIND (1985), Dori - ein Computerspiel. *Geographie Heute* 27, pp. 43-46.
- KUNZ, G.C. & F. SCHOTT (1987), *Intelligente Tutorielle Systeme*. Göttingen: Hogrefe.
- LEUTNER, D. (1988), Computersimulierte dynamische Systeme, Wissenserwerb unter verschiedenen Lehrmethoden und Sozialformen des Unterrichts. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 20, pp. 338-355.
- LEUTNER, D. (1989), Implementation und experimentelle Evaluation im computersimulierten Planspiel "Hunger in Nordafrika". *Geographiedidaktische Forschungen* 18, pp. 81-109.
- LEUTNER, D. (1990), *Simulation und Modellbildung*. Tübingen: Deutsches Institut für Fernstudien (DIF) ed., *Lehren und Lernen mit dem Computer*, vol. 1: Computereinsatz im Fachunterricht, pp. 22-52.
- LEUTNER, D. (1991), *Adaptive Lehrsysteme (Habilitationsschrift)*, Philosophische Fakultät der RWTH Aachen.
- LEUTNER, D. & R. KRETZSCHMAR (1988), Veranschaulichung und Aktivierung: Überraschende Effekte zweier didaktischer Prinzipien. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 20, pp. 263-278.
- LEUTNER, D. & H. SCHRETTENBRUNNER (1989), Entdeckendes Lernen in komplexen Realitätsbereichen, *Unterrichtswissenschaft* 17, pp. 327-341.
- MANDL, H. & P.M. FISCHER, eds., (1985), *Lernen im Dialog mit dem Computer*. München: Urban & Schwarzenberg.
- MANDL, H. & A. LESGOLD, eds., (1988), *Learning issues for intelligent tutoring system*. New York: Springer.
- NIEMZ, G., ed., (1989), *Das neue Bild des Geographieunterrichts*, Frankfurter Beiträge zur Didaktik der Geographie vol. 11.
- NUSSBAUM, A. & D. LEUTNER (1986), Die Auswirkung der Schwierigkeit textbegleitender Fragen auf die Lernleistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 18, pp. 230-244.

- RHEINBERG, F. (1985), Motivationsanalysen zur Interaktion mit Computern. In: H. Mandl & P.M. Fischer, eds., *Lernen im Dialog mit dem Computer*, pp. 83-105. München, Wien, Baltimore: Urban & Schwarzenberg.
- SCHILKE, K., H. SOENNICHSEN & R. LANGHEINE (1984), Verhaltensbeobachtungen an Schülern während eines Simulationsspiels über Landwirtschaft in einem Trockengebiet. In: R. Hedewig & L. Staack, eds., *Biologieunterricht in der Diskussion*. Köln: Aulis, pp. 208-231.
- SCHRETENBRUNNER, H. (1973), *Multi-Medien-Paket Stadtsanierung*, *Der Erdkundeunterricht* vol. 17.
- SCHRETENBRUNNER, H., ed., (1976), *Quantitative Didaktik der Geographie - Teil 1*, *Der Erdkundeunterricht* vol. 24.
- SCHRETENBRUNNER, H., ed., (1989), *Software für den Geographieunterricht*, *Geographiedidaktische Forschungen* vol. 18. Lüneburg: Hochschulverband für Geographie und ihre Didaktik (revised edition: Nürnberg 1991).
- SCHWARZER, R., ed., (1986), *Skalen zur Befindlichkeit und Persönlichkeit*. Berlin: Institut für Psychologie der FU, Forschungsbericht 5.
- SKINNER, B.F. (1958), *Teaching Machines*. Science 128, pp. 696-977.
- SNOW, R.E. (1980), Aptitude processes. In: R.E. Snow, P.A. Frederico & W.E. Montagne, eds., *Aptitude, Learning and Instruction* vol.1, Hillsdale N.J.: Erlbaum, pp. 27-63.
- SNOW, R.E. & D.F. LOHMAN (1984), Towards a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology* 76, pp. 347-376.
- TENNYSON, R.D. & D.L. CHRISTENSEN (1988), MAIS: An intelligent learning system. In: D.H. Jonassen, ed., *Instructional designs for microcomputer courseware*. Hillsdale N.J.: Erlbaum, pp. 247-274.
- WAGER, W. & R.M. GAGNÉ (1988), Designing computer-aided instruction. In: D.H. Jonassen, ed., *Instructional designs for microcomputer courseware*, Hillsdale N.J.: Erlbaum, pp. 35-60.
- WALFORD, R. (1972), *Lernspiele im Erdkundeunterricht* (Translation from: *Games in Geography*, Longman London 1969). *Der Erdkundeunterricht* vol. 14.

3 SOME PSYCHOLOGICAL GUIDELINES FOR DESIGNING INSTRUCTIONAL COMPUTER SIMULATIONS

Detlev Leutner and Beate Schreiber

3.1 Introduction

Computer simulation games like "Hunger in the Sahel" meet a lot of those criteria which - according to the work of Lepper and Malone (Malone & Lepper 1987; Lepper & Malone 1987) - are expected to make learning fun. The two most relevant criteria are the following: First, the learning situation requires the students to be active without having to face serious consequences in case of non-optimal decisions. Second, playing the game successfully is not a trivial task so that the learning situation is challenging.

On the other hand, a series of instructional experiments has shown that fun is definitely not a sufficient condition for excellent learning results, i.e. learning results in terms of the amount of knowledge about the simulated domain students extract from the simulation (Leutner 1988a; Leutner & Kretschmar 1988; Leutner 1989a,b; Leutner & Schrettenbrunner 1989; Leutner 1990; Leutner 1991; Schrettenbrunner 1992, this volume): Computer simulation games should be enhanced with teaching functions which make sure that students not only learn how to play the game but also learn something which is relevant beyond the scope of the specific game, i.e. knowledge about the simulated domain the game is taken from.

Cognitive learning from computer simulation games can only take place, if the student is able to draw correct conclusions by observing the simulated system's behavior. Usually, the network of functional relations between system variables is more or less hidden so that it is not an easy task for the student to predict the system's behavior. Furthermore, with the student playing an active role within the system, the system's behavior depends on the student's decisions and activities in controlling the system, which again complicates the learning situation. All in all, the learning situation can be described to be a complex problem situation which is characterized by missing transparency.

German psychological problem solving research of the 1980s indicated that human beings have only limited capabilities to control such complex systems (see Funke, 1985, 1986). The control performance, however, can be increased if the complexity is reduced by additional information, giving deeper insight into the functional relations between system variables. In order to be effective, however, this informa-

tion has to be relevant. Moreover, it needs to be presented appropriately in terms of amount, time, and - last not least - in terms of detail (for experimental results on the optimal level of detail of information see Leutner, 1989a; Leutner 1991).

In a series of three instructional experiments with the computer simulation game "Hunger in the Sahel" it was demonstrated that on-line adaptive advice (which is presented in situations where it seems to be necessary) and non adaptive background information (which can be looked up when desired) are powerful tools for reducing the system's complexity and thereby increasing students' cognitive learning outcomes (Leutner 1989b; Leutner & Schrettenbrunner 1989; Leutner 1990; Leutner 1991; Schrettenbrunner 1992, this volume). This is in perfect accordance with theories of teaching (e.g. Klauer 1985), theories of discovery and guided discovery learning (Nussbaum 1984; Nussbaum & Leutner 1986; Leutner 1988b) as well as cognitive psychological theories like ACT* (Anderson 1983, 1987a,b) and REPAIR-theory (Brown & VanLehn 1980).

From these experiments, however, two questions remain open:

- (1) What is more effective with regard to cognitive learning outcomes: to implement non adaptive background information on several screen pages within the computer simulation program (as realized in the first and second experiment) or to provide them on a sheet of paper (as realized in the third experiment)?
- (2) What about the effect of implementing powerful learning aids on students differing in level of inductive thinking ability: Will both, the able and the less able students profit or will one group be hindered? In other words: Are there main effects of inductive thinking ability and of learning aids, or is there an aptitude-treatment-interaction effect?

3.1.1 Background information on screen versus on paper

First of all, it should be noted that the presentation of background information proved to be an effective learning aid in previous experiments. From psychological information processing theories, however, it can be expected that presenting this information on a sheet of paper should be more effective than implementing the same information as several screen pages within the computer simulation program. The line of reasoning is as follows: If background information is provided on a sheet of paper, it is quickly accessible, at a glance, so to say. Since students often use this opportunity, the information is processed repeatedly and there is a high probability that it will be transferred from working memory into long-term memory (Atkinson & Shiffrin 1968; Nussbaum & Leutner 1986).

On the other hand, if the information is only available on separate screens within the program, the student has to realize that there is a special key on the keyboard which leads to background information. He then has to store the current system status and his action plans on a stack in working memory, leave the simulation environment, scroll through several screen pages, read and process the background information, return to the simulation environment and load back again the current system status and his plans from the stack in working memory. First, this procedure will take more effort and, presumably, will not be used very often so that the information will not be overlearned (Atkinson & Shiffrin 1968). Second, there will be a high load on memory capacity so that pieces of information might vanish (Miller 1956; Schumacher 1991). In conclusion, it is expected that the information will be transferred to long-term memory with a lower probability.

3.1.2 Inductive thinking, learning aids, and learning outcome

One of the results of German psychological problem solving research with complex computer simulation systems seems to be that problem solving performance and psychometric intelligence are correlated only around zero (Putz-Osterloh 1981; Dörner & Kreuzig 1983; Dörner et al. 1983). However, the correlation seems to increase with enhanced system transparency (Putz-Osterloh & Lürer 1981; Hesse 1982; Hussey & Granzow 1987; Hussey 1989; Hörmann & Thomas 1989). The same effect was observed in our own previous experiments with "Hunger in the Sahel", especially with regard to the correlation between inductive thinking ability and the amount of knowledge acquisition (see Leutner, 1991, p. 200): For example, in the first experiment (Leutner 1989b) the correlation was increased from $-.11$ to $.70$ by implementing adaptive advice into the program without having the students to work through a tutorial before exploring the simulation. On the other hand, however, in the third experiment the correlation was only slightly increased from $-.24$ to $.13$ when background information was absent and was decreased from $.27$ to $.06$ when background information was present.

It should be noted that these changes in correlations are sometimes quite small and far from being significant, but, nevertheless, they are converging over several experiments showing the following trend: The implementation of learning aids into complex computer simulations like "Hunger in the Sahel" seems to increase the correlation between inductive thinking ability and cognitive learning outcomes up to a point where even more learning aids will decrease the correlation. What does that mean in terms of learning opportunities for able and less able students? It can be hypothesised (see Leutner, 1991, p. 201) that in a non-transparent, complex simulation cognitive learning is purely at random, which is reflected in low outcome means and zero correlation between intelligence and outcome. If only little aid is given, then only the able students will successfully draw correct conclusions within the simulated system and thereby extract the relevant information to be learned: The overall outcome will be higher, but only because of the able students, and there will be a positive correlation between intelligence and outcome. With massive aid the less able students will follow in drawing correct conclusions without the able students benefitting any more: The overall outcome will increase again, but now only due to the less able students, and the correlation between intelligence and outcome will decrease.

If this hypothesis is correct and if one is not only interested in maximizing cognitive learning outcomes for able students, then the didactic consequence would be to use highly transparent computer simulation programs in any case. On the other hand, it has to be assured that there are no negative effects of such an easy simulation on able students, which might be expected from research in guided discovery learning (see Klauer 1985; Lohman 1986; Clark 1990a,b).

Summarizing up to this point, there are some relevant open questions with regard to the effects of student and system characteristics on the amount of knowledge which can be acquired in exploring an instructional computer simulation. An experiment was conducted to investigate, on the student's side, the effect of inductive reasoning ability and, on the system's side, on-line adaptive advice and non-adaptive background information.

3.2 Hypotheses

1. Adaptive advice with regard to system states and action alternatives will increase students' amount of knowledge acquisition.
2. Qualitative background information on system variables will increase students' amount of knowledge acquisition.
3. Qualitative background information on system variables, which is permanently available and which is easily accessible, will increase students' amount of knowledge acquisition more than the same information which can only be looked up by scrolling through several screen pages.
4. Given a computer simulation with some amount of transparency and given a positive linear correlation between students' inductive reasoning ability and students' amount of knowledge acquisition, this correlation will decrease when the system's transparency will further be increased by adaptive advice on system states and action alternatives as well as by qualitative background information on system variables.

Hypotheses (1) and (2) are system transparency hypotheses, (3) is an ease-of-accessability hypothesis, and (4) is an aptitude-treatment-interaction (ATI) hypothesis.

3.3 Methodological design

Experimental Design and variables: The experimental treatments corresponded to a 2x3-factorial design. The first factor was "adaptive advice" (levels A0: advice absent; A1: advice present). The second factor was "Qualitative Background Information" (levels I0: information absent; I1: information provided on two computer screens; I2: information provided on one sheet of paper). The dependent variable was the student's number of correct answers in a criterion-referenced posttest. Covariables were four cognitive aptitudes (pretest-knowledge, verbal fluency, inductive reasoning, speed of closure). Pretest- and posttest-knowledge were measured on a 27-item, criterion-referenced multiple-choice test developed by Leutner (1989b). Verbal fluency, inductive reasoning, and speed of closure were measured on subscales of an intelligence test (scales 1+2, 3, 12, respectively, of the LPS; Horn, 1962). Inductive reasoning was measured in order to test hypothesis no. 4 listed above. Verbal fluency was measured in order to control one of those posttest predictors which - in previous experiments - turned out to be most effective. Speed of closure was included because this ability might have an additional effect on student's opportunity to find and to process the relevant information.

Subjects: Subjects were 102 students from the Aachen University of Technology (engineering sciences, electronics, mathematics, physics, and business), 15 students being female and 87 students being male. The range of age was 19 to 30 years. Students were assigned to one of the six experimental treatments at random, 17 for each treatment. Because one student showed an obviously negative response set in all of the cognitive variables, indicated by being a clear outlier in the

distributions of data values, this student was eliminated. In order to balance the experimental design (equal cell frequencies), five other students each drawn from one of the remaining experimental treatments at random, also had to be eliminated. Therefore, each cell of the design contained 16 students so that $6 \times 16 = 96$ students remained for data analysis.

Computer Simulation Program: The simulation program "Hunger in the Sahel" was used. In order to increase the systems transparency, the original Schrettenbrunner program was changed in such a way that changes in program variables and the effects of student decisions became more obvious: For example, if the student decides to plant trees he will see in each simulated year that the soil is regenerated to a certain extent. This was obtained by changing the growth parameter from 5 cm in the original program (hardly visible on the computer screen) to 30 cm. Furthermore, the start situation was changed in such a way that the student will encounter immediately, after the first simulated year with thunderstorms, the dramatic effects of erosion.

According to the experimental design, the program was manipulated as follows:

- A0: While exploring the simulated system, the student receives no adaptive advice.
- A1: While exploring the simulated system, the student receives adaptive advice. Advice is provided, if the system diagnoses from the student behavior and/or the system status that the student lacks specific information to overcome or to cope with critical situations. A total of 34 different messages is implemented.
- I0: While exploring the simulated system, the student can only look up, on one computer screen page, numerical information on system variables in order to calculate costs and benefits of actions.
- I1: While exploring the simulated system, the student can look up three different computer screen pages with information on system variables and alternative actions: one page with numerical information in order to calculate costs and benefits of actions and two pages with qualitative background information. Compared to previous experiments, this background information is more elaborate and therefore extended from one to two screen pages.
- I2: While exploring the simulated system, the student can permanently look up numerical and background information on one sheet of paper positioned near the keyboard. The sheet of paper simply contains a hardcopy of the three computer screen pages implemented in the program version of condition I1.

The experimental manipulations were combined to obtain $2 \times 3 = 6$ different program versions according to table 3.1.

Experimental Procedure: First, the cognitive aptitudes were measured. Second, the computer simulation program was introduced to the students. Third, during a period of 45 minutes, students explored the simulation (exploration phase). This phase contained the experimental treatment. Fourth, students worked through the criterion-referenced posttest. Fifth, students worked again with a modified computer simulation (test phase). The results of this test phase, however, will be reported elsewhere. Students were assigned to one of 9 personal computers at random. The 6 different treatments were assigned to the 9 computers at random, too. Students

worked individually. While working on the pretest and the posttest, students were asked not to guess the correct answer. Details of the experimental design, the software modifications and the procedure are given in Schreiber (1991).

Table 3.1 Versions of the computer simulation "Hunger in the Sahel" and no. of subjects N within each cell of the experimental design

		Information on system variables		
		Numerical information only	Numerical and qualitative background information	
		On one computer screen page I0	On three computer screen pages I1	On one sheet of paper I2
On line adaptive advice	absent A0	A0 I0	A0 I1	A0 I2
		N=16	N=16	N=16
	present A1	A1 I0	A1 I1	A1 I2
		N=16	N=16	N=16

3.4 Results

As a first step, the overall mean of the pretest was calculated. The value is 16.0, which is - presumably an effect of the instruction not to guess the correct answer - far less than the average knowledge in previous experiments with the same material. Furthermore, the pearson correlation between pretest and posttest knowledge is $r = 0.01$ and the correlation between speed of closure and posttest knowledge is $r = 0.19$, both correlations being low and quite constant across the six experimental conditions (see table 3.4). Therefore, pretest knowledge and spatial ability were excluded from further analyses.

In a second step, a general linear model was specified for the dependent variable posttest knowledge, including main and interaction effects of the experimental treatments as well as the two remaining cognitive abilities (verbal fluency, inductive reasoning) as covariates. In order to test the homogeneity of regression slopes, aptitude-treatment-interaction (ATI-) effects were specified. Only those ATI-effects were kept within the model which indicated non-homogeneity. The remaining model and the results of analysis of covariance are given in table 3.2. The statistical alpha-error was set at 5% for all tests of significance.

Note that all effects in which the 3-level factor "background information" is involved are tested by two planned comparisons according to the hypotheses listed above: The first comparison contrasts I0 (without any background information at all) against I1 and I2 simultaneously (with background information). The second comparison contrasts the two ways in which background information is made available, I1 (on

screen pages) against I2 (on a sheet of paper). Also note, that each F-test (excluding the test of the overall effect of inductive reasoning) deals with a one-sided hypothesis about two means. Because the F-statistics reported in table 3.2 can be transformed into the t-statistic (t is the square root of F, see any basic statistics textbook), the one-sided p-values of the corresponding t-statistic can be calculated from table 3.2 by dividing those p-values by 2.

Table 3.2 Analysis of covariance (ancova) of posttest knowledge ^{a)}

Source		SS	df	MS	F	p	R ²
Cognitive Aptitudes							
- Verbal Fluency	F	22.2	1	22.2	7.311	0.008	0.063
- Inductive Reasoning	R	1.6	1	1.6	< 1	-	-
Adaptive Advice							
	A	22.0	1	22.0	7.263	0.008	0.062
Inductive Reasoning x Advice							
	RxA	18.6	1	18.6	6.135	0.015	0.053
Background Information							
- Absent vs. Present	I1	10.3	1	10.3	3.404	0.069	0.029
- Program vs. Paper	I2	4.9	1	4.9	1.608	0.208	-
Inductive Reasoning x Information							
- Info Absent vs. Present	RxI1	10.1	1	10.1	3.344	0.071	0.029
- Info Program vs. Paper	RxI2	1.9	1	1.9	< 1	-	-
Advice x Information							
- Info Absent vs. Present	AxI1	2.2	1	2.2	< 1	-	-
- Info Program vs. Paper	AxI2	1.5	1	1.5	< 1	-	-
Residual		257.9	85	3.0			

a) Each F-test deals with an one-sided hypothesis concerning the difference between two means or between two linear regression slopes. Therefore, the two-sided p-values have to be divided by 2 in order to obtain the proper significance levels.

As can be seen in table 3.2, there are five significant effects: (1) verbal fluency (F), (2) adaptive advice (A), (3) inductive reasoning x adaptive advice (RxA), (4) qualitative background information absent vs. present (I1), (5) inductive reasoning x qualitative background information absent versus present (RxI1). Inductive reasoning (R) does not have an overall effect, nor does the interaction of the experimental treatments (AxI1, AxI2) and the difference between providing the background information on several computer screen pages versus on one sheet of paper (I2, RxI2).

Figure 3.1 and table 3.3 indicate that the main effects of the experimental treatments confirm two of the hypotheses listed above: Advice is better than no advice, background information is better than no background information. But the benefit

of providing background information on a sheet of paper is so small, that the difference in knowledge acquisition is not significant (Table 3.2). Figures 3.2 and 3.3 as well as table 3.4 indicate that the two significant ATI-effects also confirm the hypothesis listed above: The linear correlation between inductive reasoning and knowledge acquisition decreases with increased system transparency (induced by adaptive advice or by qualitative background information). Note that the curved lines in the figures indicate the 75% confidence limits for the straight regression lines. Also note that the confidence intervals of the two regressions do not overlap in the lower left part of both plots. This means that the differences in posttest achievement are of practical significance only for the less able but not for the able students.

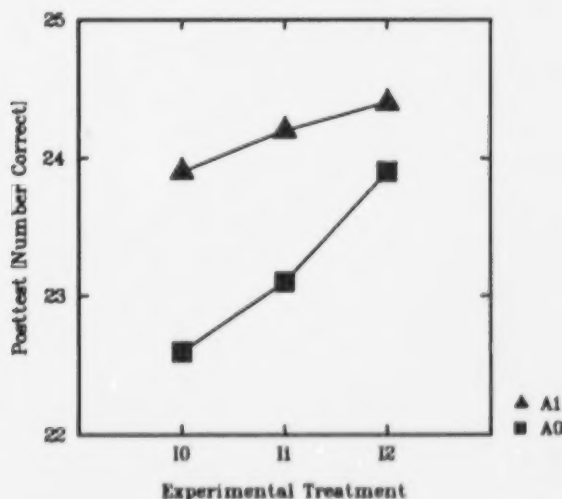


Figure 3.1 Mean posttest knowledge (A0: without, A1: with adaptive advice, 10: without qualitative background information, 11: information provided on two computer screen pages, 12: information provided on a sheet of paper positioned near the keyboard)

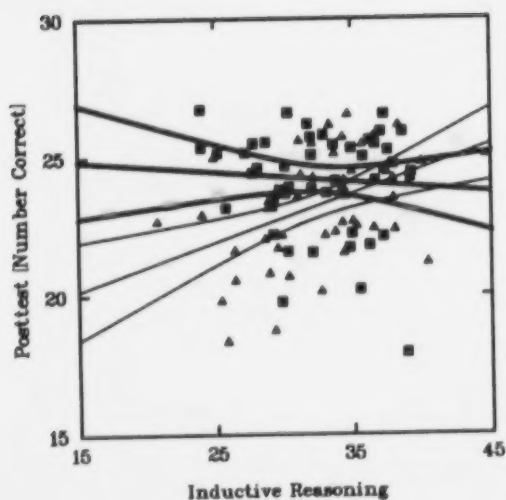


Figure 3.2 Differential regression of posttest knowledge on inductive reasoning (ATI-effect; thin lines and triangle symbols: without, thick lines and rectangle symbols: with adaptive advice)

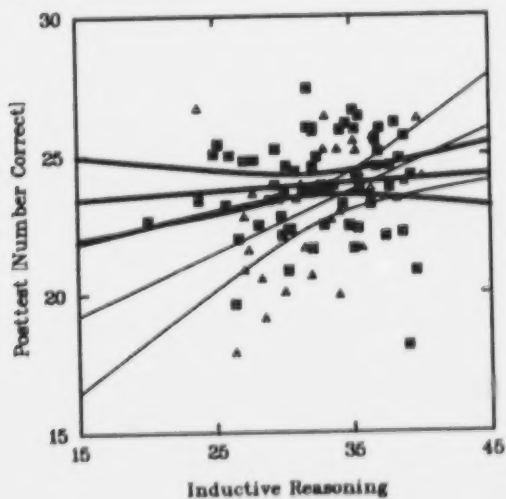


Figure 3.3 Differential regression of posttest knowledge on inductive reasoning (ATI-effect; thin lines and triangle symbols: without, thick lines and rectangle symbols: with qualitative background information)

Table 3.3 Means and standard deviations (SD) of cognitive variables

		Total	Adaptive advice		Background information ^{a)}			Advice absent A0			Advice present A1		
			A0	A1	I0	I1	I2	I0	I1	I2	I0	I1	I2
Verbal	Mean	54.2	54.8	53.7	55.1	53.8	53.9	56.0	53.4	53.1	54.1	52.1	54.8
fluency	SD	10.0	9.2	10.7	10.1	10.6	9.4	9.7	9.2	8.9	10.8	11.9	10.0
Inductive	Mean	32.6	32.3	32.8	33.0	31.8	33.0	32.4	31.8	32.8	33.5	31.8	33.3
reasoning	SD	4.3	4.5	4.1	4.2	4.7	4.0	4.4	5.2	4.0	4.0	4.3	4.3
Speed of	Mean	27.8	27.9	27.6	28.2	26.8	28.3	28.7	26.3	28.8	27.7	27.4	27.8
closure	SD	4.3	4.5	4.1	4.0	4.9	3.9	3.4	5.9	3.7	4.6	3.8	4.1
Pretest	Mean	16.0	16.5	15.5	15.7	16.1	16.3	16.9	16.1	16.6	14.5	16.1	15.9
knowledge	SD	4.1	3.7	4.5x	4.8	3.7	3.9	4.3	3.6	3.4	5.0	4.0	4.5
Posttest	Mean	23.7	23.2	24.2	23.3	23.7	24.2	22.6	23.1	23.9	23.9	24.2	24.4
knowledge	SD	2.0	1.9	1.9	2.2	2.0	1.5	2.3	1.7	1.7	2.0	2.2	1.4
N of subjects		96	48	48	32	32	32	16	16	16	16	16	16

- a) I0: Qualitative Background Information absent;
 I1: Qualitative Background Information available within the program on two separate screen pages;
 I2: Qualitative Background Information available on one sheet of paper.

Table 3.4 Pearson correlations (*r*) between cognitive aptitudes and posttest knowledge ^{a)}

	Total	Adaptive advice		Background information			Advice absent A0			Advice present A1		
		A0	A1	I0	I1	I2	I0	I1	I2	I0	I1	I2
Verbal Fluency	28	.47	.15	.24	.34	.32	.60	.57	.44	-.04	.30	.18
Inductive Reasoning	18	.41	-.08	.42	-.05	.20	.73	.15	.32	-.01	-.24	.07
Speed of Closure	19	.21	.21	.12	.27	.19	.08	.39	.14	.25	.12	.30
Pretest Knowledge	.01	.10	-.01	-.14	.14	-.09	.11	.10	.15	-.35	.39	-.26
N of Subjects	96	48	48	32	32	32	16	16	16	16	16	16

- a) Decimal points omitted

3.5 Summary and discussion

In the introduction it was argued that working with computer simulations like "Hunger in the Sahel" should be fun. And indeed, our own instructional experiments (i.e. the present experiment and previous experiments) with students aged 12 to 30, as well as Schrettenbrunner's experiments and also experiences of teachers in standard classroom settings indicated that this expectation is true: Students of a 7th class (about 13 years old), for example, spontaneously commented s.th. like "Super!", "That was great!" or "That was great fun!", as Reichwein & Rötzel (1990, p. 207) report.

Having fun is an important but, obviously, not the only goal of successful instruction: Besides fun there should be a cognitive learning outcome in terms of knowledge about concepts, facts, rules and the effects of different actions within the simulated domain beyond the framework of the specific computer simulation.

The results of previous experiments indicate that there are some restrictions concerning the amount of knowledge students are able to acquire by exploring complex systems like "Hunger in the Sahel": The critical point is that students have to make inferences from observing the system's behavior and the system's reactions on students' actions. If these inferences are wrong, no meaningful learning can take place. Therefore, computer simulation programs should be designed in such a way that students are guided to draw correct conclusions from observing changes in system variables.

The experiment reported above was conducted in order to investigate the effects of on-line adaptive advice and non-adaptive background information on the amount of learning with students differing in inductive reasoning ability. Compared to previous experiments, the computer simulation "Hunger in the Sahel" was modified in such a way that there is a higher degree of system transparency which should be effective without any additional learning aid. It was predicted, that implementing adaptive advice and providing background information would increase student's amount of knowledge, which is acquired in exploring the system. Moreover, it was predicted that background information provided on a sheet of paper would be more effective than the same information available on separate computer screen pages implemented within the simulation program. Additionally, it was predicted that providing learning aids would reduce the correlation between inductive thinking ability and cognitive learning outcome.

The experimental results are, with one exception, in accordance with these predictions: Both kinds of learning aids are effective in significantly increasing the amount of knowledge acquired and decreasing the correlation between inductive reasoning ability and knowledge. But, and this is the exception, background information provided on a sheet of paper is not significantly more effective than the same information provided on several computer screen pages within the simulation program.

The results are in accordance with the psychological system transparency hypothesis: In exploring a complex computer simulated system students will acquire more knowledge, if the functional relations between system variables and the actual

and potential effects of decisions are made more obvious by on-line adaptive advice and background information. Especially the less able students will profit from this higher degree of transparency whereas the able students are not hindered. Note that this last effect is indicated by the fact that in Figures 3.2 and 3.3 the confidence intervals of the regression lines overlap for able students indicating no significant knowledge differences whereas for less able students the confidence intervals do not overlap indicating significant differences. Also note that only the inductive reasoning variable produced this kind of aptitude-treatment-interaction effect, as it had been expected. This effect was not observed for the other two aspects of intelligence, verbal fluency and speed of closure, although for verbal fluency there is a slight trend in the hypothesised direction (with regard to on-line adaptive advice, but not with regard to background information, see table 3.4).

However, the results - as they are observed in this experiment - are not in accordance with the ease-of-accessability hypothesis: Although in figure 3.1 it can be seen that background information on paper is more effective than the same information on computer screen pages, especially for the condition without adaptive advice, the differences are too small to be statistically significant. This, however, might be due to a slight ceiling effect to be observed in the experimental group with adaptive advice and with background information on paper (A112, see figure 3.1 and table 3.3): The standard deviation of this group on posttest knowledge is lowest, although the overall differences between the six variances are not significant,

$$F_{\max} = 2.3^2 / 1.4^2 = 3.44 < F_{\max, 5\% \text{ crit.}} = 4.68, df=15).$$

3.6 Conclusion

Computer simulation systems like "Hunger in the Sahel" turn out to be powerful instructional tools. Exploring such a system not only is fun but also enables the acquisition of subject matter related knowledge which extends beyond the framework of the specific simulation used. The amount of knowledge acquisition, however, is a monotone function of the system's transparency: The transparency should be high, so that even the less able students will be successful in drawing correct conclusions from observing the system's behavior and thereby extracting the relevant knowledge. If the transparency is low, then it is to be expected that - if at all - only students which are able with regard to inductive reasoning will achieve the cognitive instructional objective. On the other hand, if the transparency is high, these able students do not acquire less knowledge compared to a learning situation in which the transparency is low.

References

- ANDERSON, J.R. (1983), *The architecture of cognition*. Cambridge: Harvard University Press.
- ANDERSON, J.R. (1987a), Production systems, learning, and tutoring. In: D. Klahr, P. Langley & R. Neches, eds., *Production system models of learning and development*, pp. 437-458. Cambridge: MIT Press.
- ANDERSON, J.R. (1987b), Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review*, 94, pp. 192-210.
- ATKINSON, R.C. & R.M. SHIFFRIN (1968), Human memory: A proposed system and its control processes. In: K.W. Spence & J.T. Spence, eds., *The psychology of learning and motivation*, Vol. 2. New York: Academic Press.
- BROWN, J.S. & K. VANLEHN (1980), Repair Theory: A generative theory of bugs in procedural skills. *Cognitive Science* 4, pp. 379-426.
- CLARK, R.E. (1990a), A cognitive theory of instructional method (Paper presented at the annual meeting of the American Educational Research Association). Boston: University of Southern California.
- CLARK, R.E. (1990b), When teaching kills learning: Research on mathemathantics. In: H. Mandl, E. DeKorte, N. Bennett & H.F. Friedrich, eds., *Learning and instruction*, pp. 1-22. Oxford, New York: Pergamon Press.
- DÖRNER, D. & H.W. KREUZIG (1983), Problemlösefähigkeit und Intelligenz. *Psychologische Rundschau* 34, pp. 185-192.
- DÖRNER, D. et al. (1983), *Lohhausen*. Bern: Huber.
- FUNKE, J. (1985), Problemlösen in komplexen computersimulierten Realitätsbereichen. *Sprache & Kognition* 3, pp. 113-129.
- FUNKE, J. (1986), *Komplexes Problemlösen*. Berlin: Springer.
- HESSE, F.W. (1982), Effekte des semantischen Kontexts auf die Bearbeitung komplexer Probleme. *Zeitschrift für Experimentelle und Angewandte Psychologie* 29, pp. 62-91.
- HÖRMANN, H.J. & M. THOMAS (1989), Zum Zusammenhang zwischen Intelligenz und komplexem Problemlösen. *Sprache & Kognition* 8, pp. 23-31.
- HORN, W. (1962), *Leistungsprüfsystem LPS*. Göttingen: Hogrefe.
- HUSSEY, W. (1989), Intelligenz und komplexes Problemlösen. *Diagnostica* 35, pp. 1-16.
- HUSSEY, W. & S. GRANZOW (1987), Komplexes Problemlösen, Gedächtnis und Verarbeitungstil. *Zeitschrift für experimentelle und angewandte Psychologie* 34, pp. 212-227.
- KLAUER, K.J. (1985), Framework for a theory of teaching. *Teacher and Teacher Education* 1, pp. 5-17.
- LEPPER, M.R. & T.W. MALONE (1987), Intrinsic motivation and instructional effectiveness in computer-based education. In: R.E. Snow & M.J. Farr, eds., *Aptitude, learning, and instruction* 3, pp. 255-286. Hillsdale, N.J.: Erlbaum.
- LEUTNER, D. (1988a), Computersimulierte dynamische Systeme: Wissenserwerb unter verschiedenen Lehrmethoden und Sozialformen des Unterrichts. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 20, pp. 338-355.
- LEUTNER, D. (1988b), Entdeckenlassendes Lehren: Beitrag zu einer facettentheoretisch begründeten empirischen Lehrforschung. In: H. Schrettenbrunner & J. van Westrhenen, eds., *Empirische Forschung und Computer im Geographieunterricht*. Niederländisch-deutsches Symposium, pp. 24-43. Lüneburg: Hochschulverband für Geographie und ihre Didaktik.
- LEUTNER, D. (1989a), Angeleitetes Lernen mit Planspielen: Lernerfolg in Abhängigkeit von Persönlichkeitseigenschaften sowie Ausma und Zeitpunkt der Anleitung. *Unterrichtswissenschaft* 17, pp. 342-358.

- LEUTNER, D. (1989b), Implementation und experimentelle Evaluation von Lernhilfen im computersimulierten Planspiel "Hunger in Nordafrika". In: H. Schrettenbrunner, ed., Software für den Geographieunterricht. Geographiedidaktische Forschungen 18, pp. 81-109.
- LEUTNER, D. (1990), Simulation und Modellbildung. In: Deutsches Institut für Fernstudien, ed., Lehren und Lernen mit dem Computer Vol. 1: Computereinsatz im Fachunterricht, pp. 22-52. Tübingen: Deutsches Institut für Fernstudien (DIFF).
- LEUTNER, D. (1991), Adaptive Lehrsysteme. Aachen: Philosophische Fakultät der RWTH. (Habilitationsschrift)
- LEUTNER, D. & R. KRETZSCHMAR (1988), Veranschaulichung und Aktivierung: Überraschende Effekte zweier didaktischer Prinzipien. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 20, pp. 263-278.
- LEUTNER, D. & H. SCHRETTEBRUNNER (1989), Entdeckendes Lernen in komplexen Realitätsbereichen: Evaluation des Computer-Simulationsspiels "Hunger in Nordafrika". Unterrichtswissenschaft 17, pp. 327-341.
- LOHMAN, D.F. (1986), Predicting mathemathanic effects in the teaching of higher-order thinking skills. Educational Psychologist 21, pp. 191-208.
- MALONE, T.W. & M.R. LEPPER (1987), Making learning fun: a taxonomy of intrinsic motivations for learning. In: R.E. Snow & M.J. Farr, eds., Aptitude, learning, and instruction 3, pp. 223-254. Hillsdale, N.J.: Erlbaum.
- MILLER, G.A. (1956), The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review 63, pp. 81-97.
- NUSSBAUM, A. (1984), Entdeckendes Lernen: Probleme der Forschung und mögliche Lösungen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 16, pp. 57-76.
- NUSSBAUM, A. & D. LEUTNER (1986), Entdeckendes Lernen von Aufgabenlösungsregeln unter verschiedenen Anforderungsbedingungen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 18, pp. 153-164.
- PUTZ-OSTERLOH, W. (1981), Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg. Zeitschrift für Psychologie 189, pp. 79-100.
- PUTZ-OSTERLOH, W. & G. LÜER (1981), Über die Vorhersagbarkeit komplexer Problemlöseleistungen durch Ergebnisse in einem Intelligenztest. Zeitschrift für Experimentelle und Angewandte Psychologie 28, pp. 309-334.
- REICHWEIN, J. & J. RÖTZEL (1990), Einsatz der Computersimulation "Hunger in Nordafrika" in einer Klasse 7. Geographie und ihre Didaktik 18, pp. 198-208.
- SCHREIBER, B. (1991), Wissenserwerb mit einem computersimulierten dynamischen System. Aachen: Institut für Psychologie der RWTH. (Diplomarbeit).
- SCHRETTEBRUNNER, H. (1992), The development and evaluation of a geography computer simulation. In: H. Schrettenbrunner & J. van Westrhenen, Empirical research and geography teaching. Amsterdam: Vrije Universiteit (Netherlands Geographical Studies 142), pp. 19-46.
- SCHUMACHER, G. (1991), Vermittlung von Problemlösefertigkeit in formalen Gegenstandsbereichen. Aachen: Institut für Erziehungswissenschaft der RWTH. (Dissertation, Arbeitstitel)

4 COMPUTER-ASSISTED EVALUATION OF CURRICULUM SOFTWARE

Günter Hess

4.1 Introduction

Up to now informal and formal tests for evaluating software in the geography curriculum have mainly been carried out on paper. This has led to valid results if the person in charge of the test or others closely involved administered the tests themselves. But if the tests were let out of their sight, the tests that were returned showed considerable deficiencies. Either whole test blocks (for example, intelligence tests) were missing or pages of the test sheets were simply skipped, etc. As a result of this experience, attempts were made to develop a test system. This system and the initial experience gained are reported on in the following.

4.2 State of the art

Computer-assisted instruction in the teaching of geography has been relatively well documented. Midgley & Walker (1985) presented a first overview of the possibilities of using the computer in the teaching of geography (on the basis of the hard- and software of the time!). Empirical studies by Schrettenbrunner (1987) and Goble (1987) on individual programs were followed by publications of an overview nature in Schrettenbrunner & Van Westrhenen (1988). There are currently empirical studies on programs by Schrettenbrunner (1987), Goble (1987), Simon & Wedekind (1980), Leutner (1991). The test design of these studies involves a computer program largely evaluating tests on paper. So the computer is only used for the evaluating program itself. Although in Leutner (1991) the computer stores the pupils' activities within the program on a diskette, achievement, personality and intelligence tests and the assessment of the program by the testee take place on paper.

Such a test design always results in a number of variables having to be entered in the computer after the work with the testee, so that, for example, they can be further processed and evaluated under SPSS. In Leutner (1991, p. 117) this involves 43 variables for a learning goal oriented achievement test and 60 for the personality test. In addition, there is also information such as age, test group, scores from IQ tests and variables from the files with the results of the simulation itself, so that about 110 values have to be entered per testee. If cognitive learning goal oriented achievement tests are also used in order to determine previous knowledge before the testee works with the simulation or if this test additionally serves to determine the medium-range long-term knowledge after the work with the simulation, then the

number of variables increases to up to 200. Leutner (1991, p. 119) worked with 87 testees in his study. In his test design this meant entering around 10,000 values.

With such an amount of data it is obvious that attempts should be made to replace the tests on paper by tests with the computer.

An attempt to carry out a software evaluation exclusively with the computer has not been undertaken so far to my knowledge. This may lie in the apparent disadvantages in practice: For the study introduced in this publication by Schrettenbrunner it was necessary to reserve the computer room at various schools up to three classroom hours long, a difficult venture even during the afternoon. In addition, the fact that many school computers are only equipped with 360 K byte disk drives led to the fact that a total of five diskettes had to be prepared per testee. The time for preparing computer-assisted evaluation is probably not considerably less than the post assessment of evaluations with tests on paper.

However, the fact that computer-assisted evaluation of software is sensible is shown by the following remarks.

4.3 The concept of the overall program package

The *Supereva/Eva* program package serves for creating and evaluating tests in the field of psychology and teaching method research. The objective is to install tests already produced in paper form on the computer, to carry them out there and to evaluate them quickly with the computer.

The *Supereva/Eva* overall system consists of four major parts:

- Creation of tests (*Supereva* Part 1)
- Creation of test sequence (*Supereva* Part 2)
- Execution of tests (*Eva*)
- Evaluation of tests (*Supereva* Part 3)

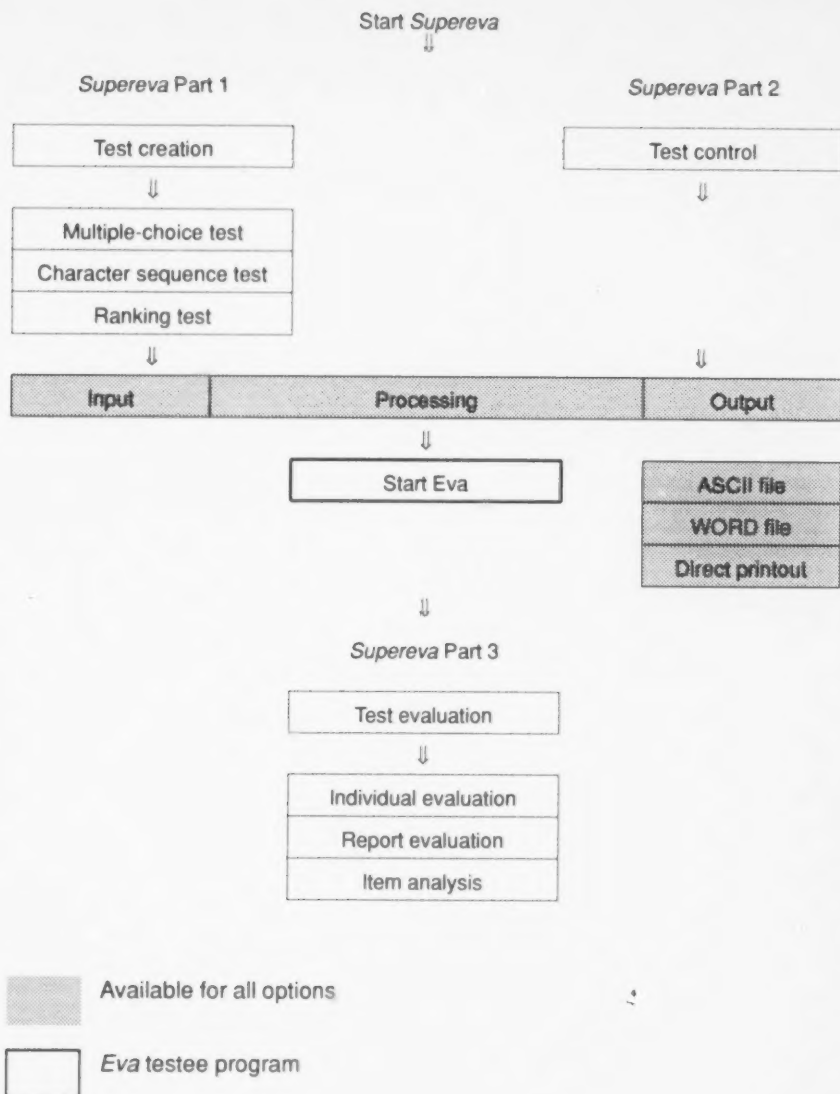


Figure 4.1 *Supereva/Eva overall system*

4.3.1 Creation of test with the *supereva* program

The first part of *Supereva* enables three test types to be created:

- a)	Multiple-choice tests with up to 5 different answer
- b)	Character sequence tests with 10 answer possibilities, which in turn can consist of one or two characters (e.g. for intelligence tests)
- c)	Ranking tests with a maximum 10-step scale

The multiple-choice tests find application within cognitive learning goal oriented tests.

30'		Test	S.	3
Bei welchem gemischten Anbau ist in Afrika Tierhaltung (Rinder, Ziegen) sinnvoll?				
Antwort 1:	Weiss ich nicht			
Antwort 2:	Hirse, Gemüse und Wald			
Antwort 3:	Hirse, Weide und Wald			
Antwort 4:	Kartoffeln, Gerste und Weide			
F1 HILFE				

Figure 4.2 Screen display of a learning goal oriented achievement test

Figure 4.2 shows such a test. A battery of 27 items for evaluating "Starvation in Africa" is documented in Leutner (1991, pp. 134-136). The following list shows a section of the achievement test Schrettenbrunner used for the study in this publication:

BEST COPY AVAILABLE

1. For which type of mixed culture is livestock production (cattle, goats) sensible in Africa?

- ☐ I don't know
- ☐ Millet, vegetables and forest
- ☐ Millet, pasture and forest
- ☐ Potatoes, barley and pasture

2. What is an example of human beings forming a landscape?

- ☐ I don't know
- ☐ Wells
- ☐ Machinery
- ☐ Terraces

3. What is erosion?

- ☐ I don't know
- ☐ The destruction of a layer of soil
- ☐ Soil with negligible slope
- ☐ Low soil fertility

4. What is the minimum that must be guaranteed by the yield of a farm business?

- ☐ I don't know
- ☐ Livestock production, schooling of the children
- ☐ Agriculture, farm machinery
- ☐ Seeds, food for the family, savings

5. What type of soil use prevents erosion?

- ☐ I don't know
- ☐ Millet
- ☐ Pasture
- ☐ Forest

6. Which measure is the best precaution against drought?

- ☐ I don't know
- ☐ Putting livestock on the range
- ☐ Making terraces
- ☐ Making wells/water holes

7. Which influence does not reduce agricultural yield?

- ☐ I don't know
- ☐ Erosion
- ☐ Low soil thickness
- ☐ Very low slope of field

8. When does the danger of erosion exist?

- ☐ I don't know
- ☐ Heavy rain on a steep slope
- ☐ High temperatures on a steep slope
- ☐ Drought on a steep slope

9. Which type of soil use is endangered most in a drought?

- ☐ I don't know
- ☐ Cotton, vegetables
- ☐ Pasture, millet
- ☐ Millet, forest

10. Which measure helps prevent erosion?

- ☐ I don't know
- ☐ Making terraces
- ☐ Making wells/water holes
- ☐ Making fenced-in pastures

However, multiple-choice tests alone do not make a study. Hence, Leutner (1991, p. 116) writes: "As can be inferred from the overview of Cronbach & Snow (1977) on the state-of-the-art for these so-called ATI effects («aptitude-treatment-interaction»), not only cognitive variables such as previous knowledge specific for the area and intelligence are to be taken into consideration as potential ATI variables, but moreover, motivation and other personality variables, too. Thus, for example in problem solving research, personality variables such as self-confidence and control competence (Funke 1985) have proven to be relevant moderator variables for the performance of the testee in using computer simulated dynamic systems".

This is the reason a possibility was incorporated to enter character sequence tests to check portions of intelligence. Such tests can consist of up to nine double characters. In addition, the program creates a question mark as the first character, standing for "I don't know" or "I don't know the word". Listed below are five items of a vocabulary test and a test to check logical thinking, respectively:

?	c	h	a	a	l	k			
?	t	a	b	e	l	s			
?	i	n	s	e	g	t			
?	g	a	s	i	n	o			
?	p	i	l	l	o	t			
(Mark the incorrect letter.)									
?	a	b	a	b	a	b	a	a	a
?	o	o	n	o	o	n	o	o	o
?	2	3	4	5	6	7	7	8	9
?	b	b	bb	b	bb	b	bb	b	bb
?	1	3	6	9	12	15	18	21	24
(Mark the incorrect character/double character.)									

As in the multiple-choice test, answers are selected using a menu bar or by clicking the answer chosen with the mouse.

As Leutner (1991, p. 128) concluded, parts of personality tests are able to explain testee performance in the simulation. Therefore, a module was incorporated with the help of which tests could also be created, in which questions could be solved using a ranking scale. This module is thus suitable for creating personality tests (figure 4.3). Questions in such tests could be:

1. I fear surprise tests.
2. When I'm confronted with a problem, I often have a number of ideas on how to handle it.
3. If I don't succeed at something immediately then I become insecure and give up.
4. My heart starts beating rapidly when teachers return assignments.
5. I am able to solve difficult problems if I make an effort.
6. I usually don't even try to solve a problem.

7. I'm afraid of exams.
8. I know exactly what I must do in order to get good grades.
9. I don't have enough patience *to deal with difficult problems endlessly*.
10. I make mistakes in tests because I'm much too afraid.

10'	Test	S.	3
-----	------	----	---

Frage 1:

Ich habe Angst davor,
dass überraschend ein Test geschrieben
wird.

stimmt nicht <-----> stimmt

☒ 1	2	3	4	5
------------------------------------	---	---	---	---





F1 HILFE

Figure 4.3 Eva testee program - question in a personality test

Part 2 enables control files to be created, which can call up a maximum of 81 tests with a pre-determined number of questions and a fixed time limit. This is supposed to prevent deficiencies as might appear when carrying out paper tests (exceeded time limit, skipped questions, etc.).

Basically it is possible to implement TenCORE programs in binary form. The possibility of calling up " *.com "- and " *.exe " files is planned. In this way it would be theoretically possible to combine a complete test design with all tests and the program used for evaluating on one diskette or the hard disk. These applications, however, are relatively unimportant as long as many school computers can only work with 360 K byte diskettes and have no hard disk. The testee program described briefly in the following, the TenCORE runtime module and a (not too simple) program go beyond the capacity of such diskettes (see 4.3.4).

The *Eva* testee program accesses the above-mentioned control files. Taking into consideration the given number of questions and time limits, it displays the questions on the screen for the testee and stores the answers.

Part 3 of *Supereva* in turn evaluates the answer files created by *Eva*. In this part records can be created from individual testee files, essentials of which can be processed statistically. The results can be displayed in graphics on the screen. Interfaces to standard software such as Word and SPSS are implemented by program parts that convert the character sequence to the Word format or to a fixed file format for SPSS. This enables data transfer. Moreover, there is a possibility of direct printout on IBM Proprinters. Adaptation to other printers is planned.

4.3.2 The *Eva* program for use by testee

The *Eva* program works to a large extent automatically, according to the information of the tester. After the testee enters last name, first name, location and age, the information on time and a number of questions appears. From this point on the testee can no longer leave the program within this test (except with CTRL+ALT+DEL or by switching off the computer). An escape possibility is to use the key combination CTRL+ALT+e; however, this should in no case be told to the testee. In tests with time limits, the time still available is displayed in the top menu bar. The selection of the answers/ranks takes place via cursor keys + ENTER/RETURN or by clicking the mouse.

30'	Test	S.	3																													
ERLÄUTERUNG: Du kannst den Balken mit den Cursortasten bewegen, Ferner stehen Dir verschiedene Funktionstasten zur Verfügung <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; padding: 5px; vertical-align: top;"> Cursortasten: <table style="width: 100%;"> <tr> <td style="text-align: center;">7</td> <td style="text-align: center;">8↑</td> <td style="text-align: center;">9</td> </tr> <tr> <td style="text-align: center;">4←</td> <td style="text-align: center;">5</td> <td style="text-align: center;">6→</td> </tr> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2↓</td> <td style="text-align: center;">3</td> </tr> <tr> <td colspan="3" style="text-align: center; padding-top: 10px;">Einfg</td> </tr> </table> </td> <td style="width: 50%; padding: 5px; vertical-align: top;"> Funktionstasten: <table style="width: 100%;"> <tr> <td>ESC</td><td>F1</td><td>F2</td><td>F3</td><td>F4</td><td>F5</td><td>F6</td><td>F7</td><td>F8</td><td>F9</td><td>F10</td> </tr> </table> </td> </tr> <tr> <td colspan="2" style="padding: 5px; vertical-align: top;"> Maussteuerung: Antwort 1 </td> <td colspan="2" style="text-align: right; vertical-align: middle;"> </td> </tr> </table>				Cursortasten: <table style="width: 100%;"> <tr> <td style="text-align: center;">7</td> <td style="text-align: center;">8↑</td> <td style="text-align: center;">9</td> </tr> <tr> <td style="text-align: center;">4←</td> <td style="text-align: center;">5</td> <td style="text-align: center;">6→</td> </tr> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2↓</td> <td style="text-align: center;">3</td> </tr> <tr> <td colspan="3" style="text-align: center; padding-top: 10px;">Einfg</td> </tr> </table>	7	8↑	9	4←	5	6→	1	2↓	3	Einfg			Funktionstasten: <table style="width: 100%;"> <tr> <td>ESC</td><td>F1</td><td>F2</td><td>F3</td><td>F4</td><td>F5</td><td>F6</td><td>F7</td><td>F8</td><td>F9</td><td>F10</td> </tr> </table>	ESC	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Maussteuerung: Antwort 1			
Cursortasten: <table style="width: 100%;"> <tr> <td style="text-align: center;">7</td> <td style="text-align: center;">8↑</td> <td style="text-align: center;">9</td> </tr> <tr> <td style="text-align: center;">4←</td> <td style="text-align: center;">5</td> <td style="text-align: center;">6→</td> </tr> <tr> <td style="text-align: center;">1</td> <td style="text-align: center;">2↓</td> <td style="text-align: center;">3</td> </tr> <tr> <td colspan="3" style="text-align: center; padding-top: 10px;">Einfg</td> </tr> </table>	7	8↑	9	4←	5	6→	1	2↓	3	Einfg			Funktionstasten: <table style="width: 100%;"> <tr> <td>ESC</td><td>F1</td><td>F2</td><td>F3</td><td>F4</td><td>F5</td><td>F6</td><td>F7</td><td>F8</td><td>F9</td><td>F10</td> </tr> </table>	ESC	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10								
7	8↑	9																														
4←	5	6→																														
1	2↓	3																														
Einfg																																
ESC	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10																						
Maussteuerung: Antwort 1																																
F1 HILFE 																																

Figure 4.4 *Eva* testee program - on-line Help window

After every entry, the program automatically calls up the next question. At the end of such a test, the testee has the choice of leaving the program or continuing with the next test. If the testee quits the test here, the program automatically calls up the next test the next time it is started up. It is not possible to leave out or to repeat a test. Likewise the program prevents the exceeding of time limits or the skipping of questions. If the testee deletes files on his diskette, the program begins with the first test as a rule.

4.3.3 Hardware specific limitations

German Stage I and Stage II secondary schools are equipped to a great extent with computers which represent a minimal PC standard: monochrome graphics card, 8086 processor and one or two 360 K byte disk drives without a hard disk. This means that from the very beginning the question section of the test program cannot be larger than 360 K (including the TenCORE specific runtime module, the video adapter driver and the anticipated pupil file). Moreover, the program must be Hercules compatible and, of course, run without a mouse. It became evident quite soon that certain parts of the program of the color version created in EGA mode had to be reworked, since both in the schools and at the university EGA and EGA-Wonder cards existed which could not display a number of colors or sometimes emulated the colors so much that they could no longer be recognized on the screen. This problem involved the program for evaluating since the *Eva* test program even gets by with few colors in the EGA version. As Watson (1987, p. 106) suggested for intelligent tutoring systems: "Keep to the minimum number of colors if possible, and select ones that provide both good contrast and legibility. ... Legibility will be best achieved by selecting two colors of contrasting brightness, for instance, black on cyan, or white on blue".

4.3.4 Problems between hardware and software

There were also problems with equipment that didn't have a hard disk. This equipment reserves a portion of the memory as a virtual hard disk. Variables were stored there in such a way that they disturbed other programs. For example, the working times of the testees were calculated incorrectly even though at the beginning of such a test the elapsed working time was automatically set to zero. Only a complete reset of the computer made the programs run properly again. In the designing and programming of such a test program, the knowledge of such difficulties should lead to variables being checked at selected positions (e.g. after setting a variable to zero or before and after reading in a data record) and error messages output, if necessary.

4.4 Initial observations on user interfaces and content design

4.4.1 User interface for the *Supereva* program

The practical work with the current version of the *Supereva* program showed that evaluation programs need to have more than just the functions of common databases. In the design of such a program the consequence is that a user interface with push buttons always present (fields that can be activated with the mouse) is not possible. A double bottom menu bar as shown in figure 4.5 represents a temporary solution in any case. The logical consequence is to change the menu logic to a pull-down menu, the greater portion of functions of which is only visible when the main menu item is called up in the top menu bar (see programs like MS Windows, MS Word 5.5, etc.).

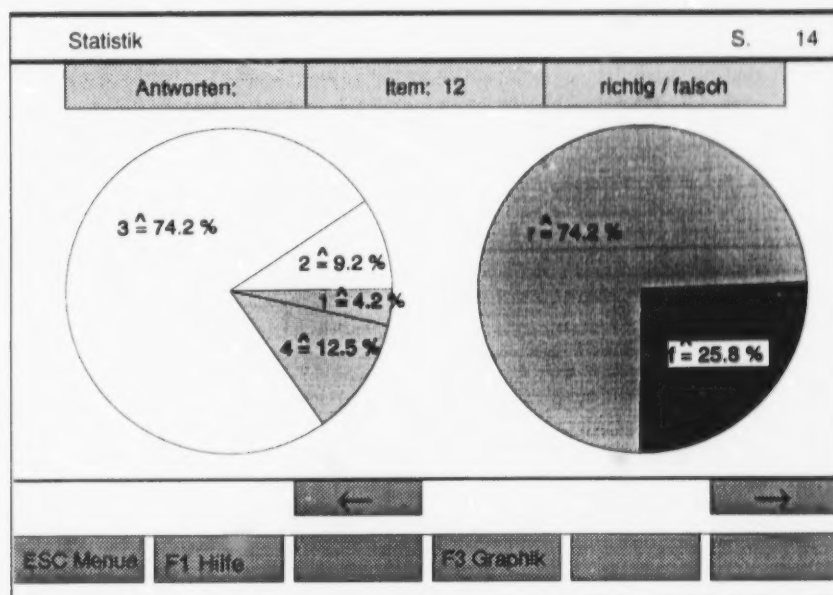


Figure 4.5 Item analysis - graphics display (here pie-chart)

4.4.2 The user interface for the *Eva* test module

The program described above is kept deliberately simple in its user interface. It was mastered quickly by the pupils. The idea of paging with the cursor right key or closing a window (figure 4.4) caused the pupils difficulties. However, the pupils who were familiar with the computer asked immediately about the possibilities of program control with the mouse when the mouse arrow appeared and did control the program

to a large extent with the mouse. The conclusion for the program designer should be that "simple" keys like ENTER/RETURN or the space bar are more obvious for paging in the program than control via cursor keys. The pupils' handling of the computer also showed the various prerequisites with which pupils approach the computer in the first stage of the German secondary school. This confirms what Egyhazy & Hutson (1985, p. 307) support: "Output from the computer to the child should be congruent with the child's current level of knowledge, logical ability, interests, and preference for various modes".

4.4.3 Design of test module contents

As already mentioned, the program package allows numerous possibilities for test design. It was done without the implementation of graphics, since the direct implementation of vector oriented graphics in the current program language is not possible by an end user who is not the program designer. The implementation of pixel graphics would be easily possible but is not advisable due to the heterogeneity of graphics standards in schools.

Unfortunately, owing to programming techniques, it cannot be avoided that a test on paper does in fact differ in design from a computer controlled test with the same questions. In contrast to the test on paper, the pupil sees only one question on the screen at a time during the achievement test. He doesn't know which questions will still appear nor can he glance through the questions already answered again. The program is designed in such a way that, for example in a vocabulary test, he sees only ten items at a time - in contrast to a complete list of words on paper.

Due to testing theory observations, the first answer possibility for each question is "I don't know". The bar is placed on this answer as the default when the question is called up. An evaluation of 30 answers in an achievement test (120 testees), however, shows that the acceptance of such an alternative answer is relatively low. Only 105 of the total 3600 answers were these alternative answers (just about 3%). The minimum, related to the individual item, was 0; the maximum 8 (6.7%).

4.5 Initial results from evaluations with *Supereva*

4.5.1 Comparison of results between paper test and computer controlled test

It has shown that the results between computer controlled tests and tests on paper deviate, although in both cases - as described - the possibility of refusing an answer is incorporated:

It is evident that all of the average values in the computer test lie considerably below those in the paper test. This can be explained easily in the vocabulary test since the test design forced the pupil to answer the questions in sequence. He could not - as on paper - go from question to question at will. Thus, by thinking about one concept for a long time, he could lose so much time that the test was ended by the computer although the pupil might have known other answers. Jäger (1990, p. 109)

also points out that for a computer-assisted test generally more time is necessary than for a paper test: "The intensive confrontation with the adaptive testing leads to the conclusion that the testee needs more time than for the conventional form (Wild 1988). Thus, it belongs to the comprehensive tasks within the scope of technology assessment to make reliable statements on such secondary conditions. In our opinion these have been currently unresearched to a great extent due to other interest situations". On the other hand, in contrast to Leutner (1991), work was done with pupils of the 6th grade (in Leutner, the 8th grade), so that differing abilities in vocabulary might be explained easily by differences in age alone.

Table 4.1 Scores on paper test and computer test

	Paper test		Computer test	
	Average	Standard deviation	Average	Standard deviation
Vocabulary	41.1	8.1	35.75	9.1
Exam nerves	2.0	0.5	1.6	0.6
Self-effectiveness	2.5	0.5	2.0	0.7
General fear	1.9	0.4	1.4	0.6
Helplessness	1.7	0.5	1.3	0.4
Curiosity	2.4	0.4	2.1	0.8
Emotional lability	2.1	0.6	1.6	0.7

Source: Paper test: Leutner 1989, p. 123

What is more difficult are observations on why pupils get lower scores for personality tests no matter what kind of characteristics (e.g. curiosity, general fear) are involved. Here the test design might have played a role again, since the cursor for marking the intensity was always placed at 1 ("not true") when the question was called up and had to be moved from there. Further studies are necessary to clarify this.

4.5.2 Initial results of statistical evaluation

The main objectives of the *Supereva* program have already been described above. One objective within the evaluation routine is to incorporate standardized analyses (Lienert 1989) in German language testing psychology and thus make them comparable with Anglo-American statistics. Here, an important routine is the direct reading in SPSS of the values calculated with *Supereva* in this way. A first result was achieved in my opinion by the comparison for the first time of Cronbach's α for all α 's, when the item to be analyzed is excluded from the calculation (α if item deleted), with the selectivity analysis of Lienert (1989, p. 93). Although the basic formulae assume different initial values and are structured differently (Lienert 1989, p. 97; Bortz 1977, p. 692), the evaluation of an item analysis showed that the selectivity coefficient of Lienert and Cronbach's α correlate with - 0.99 if posttests are employed. On the basis of this almost complete correlation, the selectivity

coefficient of Lienert is preferable to Cronbach's α due to greater concreteness at least in this study. Incidentally, an observation of the correlation matrix shows that the selectivity coefficient and α values lie almost on one line, of course, but there are some values that are so far from the regression line that a simple mathematical transformability of the formulae is to be excluded.

4.6 Conclusions

The results of a study presented in this publication by Schrettenbrunner are probably, at least in Europe, among the first to have been gained exclusively with the computer in the field of the evaluation of software in the curriculum. The experience gained in this test series can be summarized as follows:

- * Although all the test computers used were DOS machines and thus completely compatible, it was apparent that equipment without a hard disk was susceptible to errors since temporary hard disks were created in the memory by the operating system. The consequence here is either a test design which makes it possible to carry out all tests with one's own computer (e.g. laptops) or to check out the programs on every computer at the test schools to eliminate the possibility of errors.
- * Assumptions on what a program should look like regarding the user interface have been presented but never tested under school conditions. It was not the objective of this study to test the user interface but to reveal the problems of the pupils. They are not familiar enough with the computer to carry out paging on the screen with the cursor keys immediately even though the pupils quite quickly learned the user guidance from the person in charge after appropriate explanation. Help windows that open immediately during incorrect operation should be incorporated in the programs.
- * Transferring standardized tests available previously as paper tests to the computer causes a number of problems. Up to now only differences in the results have been able to be determined. Attempts at explanation may be found in the pupils' lack of the overall picture of the whole test or the user guidance involved (starting position of the cursor).

References

- BAUER, H. & U. FREIBERGER (1991), Graphische Benutzeroberflächen SAA-Standard, BUS 21, pp. 27-31. Zentralstelle für den Computer im Unterricht, Augsburg.
- BORTZ, J. (1977), Lehrbuch der Statistik für Sozialwissenschaftler. Berlin: Springer.
- CRONBACH, L.J. & R.E. SNOW (1977), Aptitudes and instructional methods. New York: Irvington.
- EGYHAZY, C.J. & B. HUTSON (1985), Principles for interfaces between young children and computers. In: K. Duncan & D. Harris, eds., Computers in education, pp. 305-309. New York.

- FISCHER, M. & H. MANDL (1988), Improvement of acquisition of knowledge by informing feedback. In: H. Mandl & A. Lesgold, *Learning issues for intelligent tutoring systems*, pp. 187-241. Berlin: Springer.
- GOBLE, T. (1987), A knowledge base for an expert system in geography. In: W.A. Kent & R. Lewis, *Computer-assisted learning in the humanities and social sciences*, pp. 146-150. Oxford: Blackwell Scientific Publications.
- HEINES, J.M. (1984), *Screen design strategies for computer-assisted instruction*. Bedford: Digital Press.
- HERBIG, M. (1976), *Praxis lehrzielorientierter Tests*. Düsseldorf: Schwann.
- JAEGER, R.S. (1990), Computerdiagnostik - Ein Überblick. *Diagnostica* 36/2, pp. 96-114.
- KLAUER, K.J. (1987), *Kriteriumsorientierte Tests*. Göttingen: Verlag für Psychologie, Hogrefe.
- KUNZ, G.C. & F. SCHOTT (1987), *Intelligente Tutorielle Systeme*. Göttingen: Verlag für Psychologie, Hogrefe.
- KYLLONEN, P.C. (1991), Principles for creating a computerized test battery. *Intelligence*, pp. 1-15.
- LEUTNER, D. (1988), Entdeckenlassendes Lehren: Beitrag zu einer facettentheoretisch begründeten empirischen Lehrforschung. In: H. Schrettenbrunner & J. van Westrhennen, eds., *Empirische Forschung und Computer im Geographieunterricht*. Amsterdam: Vrije Universiteit (Netherlands Geographical Studies 62).
- LEUTNER, D. (1991), Implementation und experimentelle Evaluation von Lernhilfen im computersimulierten Planspiel "Hunger in Nordafrika". In: H. Schrettenbrunner, ed., *Software für den Geographieunterricht. Geographiedidaktische Forschungen* 18, pp. 112-137. Nürnberg.
- LIENERT, G. A. (1989), *Testaufbau und Testanalyse*. Weinheim: Psychologie Verlags Union.
- MIDGLEY, H. & D. WALKER (1985), *Microcomputers in geography teaching*. London: Hutchinson.
- SCHRETTEBRUNNER, H. (1987), Evaluation of a program on "Distance". In: W.A. Kent & R. Lewis, *Computer-assisted learning in the humanities and social sciences*, pp. 146-150. Oxford: Blackwell Scientific Publications.
- SCHRETTEBRUNNER, H. (1991), *Software für den Geographieunterricht*. Nürnberg: Hochschulverband für Geographie und ihre Didaktik.
- SCHRETTEBRUNNER, H. & J. VAN WESTRHENEN (1988), *Empirische Forschung und Computer im Geografieunterricht*. Amsterdam: Vrije Universiteit (Netherlands Geographical Studies 62).
- SIMON, H. & J. WEDEKIND (1980), Das Computer-unterstützende Planspiel Tanaland als Test- und Trainingsinstrument zum Problemlösen in komplexen Systemen. In: H. Simon, *Computer-Simulation und Modellbildung im Unterricht*. München, Wien: Oldenbourg.
- SIMON, H. (1980), *Computer-Simulation und Modellbildung im Unterricht*. München, Wien: Oldenbourg.
- WATSON, D. (1987), *Developing CAL: Computers in the curriculum*. London: Harper & Row.
- WILD, B. (1989), Neue Simulationsstudien zur Effizienz verschiedener Parameterschätz- und Itemauswahlstrategien beim "tailored-testing". In: K. D. Kubinger, ed., *Moderne Testtheorie*, pp. 163-178. Weinheim: Psychologie Verlags Union.

5 THE COGNITIVE COMPLEXITIES OF ACCESSING CURRICULUM SPECIFIC DATABASES

Rod Gerber and John Lidstone

5.1 Introduction

Databases are general purpose software tools that have been widely adopted for classroom use. The justifications for this adoption have been classified by Degl'Innocenti & Ferraris (1988) in terms of:

1. their importance in familiarising students with programs that are increasingly used in many professional activities;
2. the assumption that using databases can enhance general logic and cognitive skills; and
3. the possibility of using databases to conduct wide ranging school projects based on large quantities of data.

In the context of geography curricula at primary and secondary school level, Lidstone (1989) supports the first two of these assumptions and discusses the application of curriculum specific databases to geography teaching. The range of such databases available to students and teachers of geography and social science in Australia are described by Lidstone & Gerber (1989). In this earlier work, we suggested that while the range of curriculum specific software available offers great potential for achieving a wide range of educational aims, variations in their design and application in school contexts may result in less than optimal achievement by both teachers and students who use them.

The purpose of this paper is to consider the influence of the interfaces used in the selected software packages as students' interact with curriculum-specific databases currently available for use in schools in Australia, and to report the results of an empirical study in which two such databases were presented to groups of students of secondary age.

5.2 Factors that influence students' success with computer-assisted learning

An examination of the literature on possible causes of less than optimal achievement of educational aims when computer-based databases are used in classrooms suggests that there are three major factors that can influence success. These are:

1. characteristics of the student;
2. characteristics of the learning environment; and
3. characteristics of the programs used.

Of these three factors, research on the first two is to be found in the literature of teaching and education and for an understanding of the third, it is necessary to look to the literature of interface engineering.

5.2.1 Student characteristics

In terms of student characteristics, Delcos & Kulewicz (1986) cite a number of papers that report research on learning characteristics, metacognitive functioning, learning styles and gender preferences. They summarise this work by stating that the authors of these studies attempt to classify children based on their approaches to computer activities, determine which characteristics of learners are the best predictors of performance, and suggest which programs or instructional techniques will benefit each group. The majority of such studies have started by defining learning objectives and have then tested various programs or packages. Delcos and Kulewicz point out that, on the basis of such studies, in general there is very little evidence that computer based efforts to teach thinking skills have been successful. The present study focuses on the qualities of the program interfaces.

5.2.2 Learning environment

The learning environments in which computers have been introduced have been typified by studies by Underwood (1988) and Watson (1990).

Underwood studied the aims of teachers in requiring students to work with databases and found that these aims fell into the following series of clearly defined groupings:

- Skills and knowledge
 - Cognitive
 - Communication
 - Computer
- Strategies
 - Active learning
 - Emancipation
 - Integration
 - Motivation
- Socio-psychological
 - Classroom
 - Self esteem
- Societal issues
 - Life skills
 - Ethics

Depending on the ages of the students, these skills were seen as being developed through three levels of interrogation of the database:

1. Paging the data whereby students were asked to reproduce on the screen specified blocks of the file with no manipulations;
2. Fact seeking whereby the machine is used to manipulate the data by a simple sort routine or to identify examples from the file according to specified criteria; and
3. Relational inquiries whereby students sought relationships and patterns within the data.

Underwood found that in defining their aims, teachers were concerned primarily with uses that place a higher cognitive demand on children in terms of Bloom's taxonomy (1956), going beyond the acquisition of knowledge to evaluation and making judgments and finally drawing together disparate information into a whole in order to solve problems. The most common aims expressed by her teacher respondents related to the cognitive and communication skills. The teachers recorded fewer comments about computer skills than about higher level skills despite the fact that most transactions and outcomes were related to the efficient use of the machine.

In terms of classroom computer use, Watson cites research by Bliss (1986) which shows that the successful implementation of innovation involving computers in education depends on access to computers, teachers' attitudes to technology, confidence in the machines and time to learn how to use both hardware and software; by Hurd (1987) which shows that teachers who adopt CAL are more likely to be predisposed towards active styles of learning, including group work, and by Johnson (1987) who claims that teachers are more likely to use new technologies and innovative software if it builds on their existing experience and knowledge of similar subjects in other forms of teaching. From this she claims that the successful integration of computers depends not only on curriculum specific software and a predisposition of teachers towards innovation, but also on the appropriate environment to support such teaching.

In making this statement, Watson is referring to the physical environment in which the machines are located (classroom vs computer room) but the same caveat may also be suggested in terms of the computing environment itself. For this third factor in considering the success of CAL, the software itself must be considered.

5.2.3 Computer assisted learning

Given the fact that in school applications of computer assisted learning, the software will usually be selected by the teacher on the basis of the machine available and the "content" of the program. However, the most important difference noted by the students is likely to be the user interface. Unless the teacher guides students by specific lessons in the specifics of particular interfaces, it may be suggested that students' attitudes towards and success with any computer-assisted learning experience will be strongly influenced by their first encounter with the interface of any new piece of software.

This contention in the context of school use is supported by various writers in software engineering. Moran (1981) defines the user interface of a computer system as those aspects of the system that the user comes in contact with physically, perceptually or conceptually, while Sommerville (1982) points out that the user interface is the yardstick by which the user judges the quality of a computer system. Both Moran & Norman (1983) have emphasised that in interacting with computer systems, people form internal, mental models. Norman goes on to claim that "most people's understanding of the devices they interact with is surprisingly meagre, imprecisely specified, and full of inconsistencies, gaps and idiosyncratic quirks". The inadequacies of these mental models have been used by Sommerville (1982) to account for the characteristic responses to poor interface design identified by Eason et al.

1. Users refuse to make use of the system and after initial "testing" deem it inadequate for their needs;
2. Users learn a few commands "parrot fashion" and restrict their usage of the system to these commands. Because the users have no real understanding of the system, they cannot make use of all the facilities; and
3. Users interact with the computer system by using an intermediary who is the local system "expert".

While these responses were formulated on the basis of experiences with adult users, the responses of school students may be very similar, both in terms of their willingness and interest in using the software and their success in using the more sophisticated facilities available to them.

5.3 Types of user interface found in curriculum specific databases

The empirical study reported here is based on comparisons of two software packages available at relatively low cost to schools in Australia: One World countries Database and Volca in the Disasters: Atmospheric and Geomorphic kit.

One World is an integrated database management system with information on 178 countries each classified within 33 fields. The management system is menu driven with the first or main menu offering the options:

- * Display a country
- * Search using names
- * Analyse data
- * Editor
- * Help reference
- * Exit to system.

There is, however, nothing in this menu to tell the neophyte user that only the first three features refer to gathering data from the database. In use, the first menu option displays all the information in the database on a specified country while the second identifies the country corresponding to specified previous name, capital city, leader of government or Head of State. Only in the third option, "analyse data", is the user

faced with a further series of menus in which choices must be made for twenty categories of information (population size, percentage of area cultivated, etc.) and to specify parameters (export = oil, desert = 2%, etc.) which may be subject to conditions and/or. Having analysed the data according to the specified conditions, a further menu permits the user to:

- * List countries names only;
- * List with selected details;
- * Try different data;
- * Do a new analysis; and
- * Exit to main menu.

The Volca database is one element of the Disasters: Atmospheric and Geomorphic software package written for use with the Nexbase+ database management program. The database contains details of most of the large volcanic eruptions under eight fields including location, year and amount of damage.

Although the authors of Nexbase+ claim that the database management system is menu driven, it has many features of command driven systems. Thus, although loading the program offers the menu options: define a query, access information, statistics and OS command, instructions thenceforth must be entered as commands.

Both these systems were designed primarily for school audiences, but while One World was designed as a fully integrated management system, Nexbase+, prepared by the Angle Park computing centre (South Australia) for school use is used as the management system for the data files provided in the Volca package. In effect, however, this distinction is not apparent to users since both packages as provided to schools include the management system and database on a single disk. Similarly, both systems are designed for use with Acorn machines, although One World is also available in a MS DOS version. Once again, the presentation of both versions to a user is identical.

In summary, in terms of the kinds of geographical questions that students and teachers are encouraged to answer in their manuals, both systems were designed for similar tasks, hardware systems and user groups.

However, Brazier & Trimp (1988) followed in adopting Moran's Command Language Grammar to differentiate between the two user interfaces. Moran distinguished four levels of interaction: task level, semantic level, syntactic level and key stroke level. While the two databases were similar at the task level and the semantic level (the elements and structures employed for the execution of the retrieval and data sorting tasks), they differed at the syntactical level where the actual interaction of the user with the system takes place. In One World, a series of menus is presented and the user invited to select from each. In Volca, the specification of actions occurs directly through the execution of commands, and although the commands are printed at the foot of each screen, each command must be entered in full in response to a prompt.

5.4 The empirical study

In a comparison of two other databases intended for school use, Brazier & Trimp administered both instruction and training to their respondents before testing the relative effectiveness of the two types of database interface. Brazier & Trimp found that under these circumstances, no difference in the ease with which their respondents came to terms with the two systems. However, in view of the many different curriculum-specific packages available for schools, we suggest that it is unlikely that students will be trained in each of the systems. Certainly, in a single curriculum area such as geography it is unlikely that the levels of training and support made available in the Brazier and Trimp study would be available. Accordingly, the present study was designed to explore the ways in which students react to databases they have never before encountered, the factors which mitigate in favour of success in their endeavours and whether, in such circumstances, menu or command driven systems are more appropriate.

5.4.1 Method

A total of 53 students at a Brisbane boys' school participated in the study. They included 25 students in a year 8 geography class, with a mean age of 12.96 years and 28 students in a year 12 geography class with a mean age of 17.38 years. The students in each class agreed to participate in the study.

The study was undertaken in three parts:

1. The students completed a questionnaire which canvassed their previous experience with computers both at home and in school, with specific attention to their previous experience with databases;
2. The students completed standardised tests designed to measure spatial intelligence and verbal intelligence; and
3. The students undertook a number of tasks based on each of the two databases. In order to prevent transfer of learning favouring one particular database, each user group was randomly split in two with one half beginning with tasks based on One World while the other half began with Volcanoes.

5.4.2 Instruments

a) The student questionnaire

The student questionnaire gathered information on attitudes to computers, students' level of experience with computers, their home ownership of a computer, the use made of a home computer, the extent to which computers had been used within their school curricula and any specific experience of databases.

b) The standardised tests

Two standardised tests were administered to the student participants in order to obtain measures of their spatial abilities (Space Test from the Maths Profile Series, ACER 1978) and general verbal abilities (Otis Lennon School Ability Test: Advanced Form R, Australian Adaptation, ACER 1987). These two abilities were seen as significant to the present study since:

- i) both databases require students to relate to specific and complex verbal instructions, and
- ii) menu driven systems require the examination of whole screens of information as well as assuming users possess a three dimension spatial analogy of nested menus.

c) The tasks required of each student on each database

This aspect of the data collection was undertaken in two adjacent computer laboratories in the students' school. The participants were paired according to their ages, scores on the standardised tests and levels of prior experience with computers and each pair was allocated to a computer already loaded with one of the database programs. The computer screen showed the "title page" of the program with the first instruction on how to proceed visible. The manual/student materials for each program was available on the desk beside the machine. Each pair of students was then given a booklet in which was printed a series of tasks, and the nature of each task was explained.

The tasks were structured in such a way that students were encouraged:

- i) to experiment with the database management system;
- ii) to initiate their own questions and seek answers from the database; and
- iii) to seek answers to a series of five specific questions which reflect both the three levels of database search: paging; fact seeking and relational queries; and the types of questions that students are encouraged to formulate in inquiry based geography syllabuses.

Thus, the five questions specified for the One World database were:

1. What are the latitude and longitude of the capital of Norway?
2. For which country is Alesandro Pertini the Head of State?
3. How many countries are there in South America?
4. How many countries, where over 80% of the land area is desert, are oil exporters?
5. What conclusions can you draw from the data in the database about the type of government of oil exporting countries?

In response to all three types of task, the participants were asked to record their interactions with the computer and the responses received under the headings:

What you did?

What happened on the screen?

What information did you find out?

What are your comments on what you did?

5.5 Analysis of the data

The analysis in this study focussed on two areas:

1. the establishment of any relationships amongst the dependant variables to determine possible relationships between success and efficiency in using menu driven and command driven databases, and
2. an investigation of linkages between the dependant variables and a set of potentially relevant independent variables associated with the use of databases.

While the independent variable data from the questionnaires was collected in numerical form and was therefore immediately amenable to statistical analysis, the data collected from students' responses to the tasks based on use of the databases required further processing. The amount of detail recorded by participants varied considerably, but in most cases it was possible for the researchers to reconstruct the processes undergone in order to achieve a final answer to the questions, whether these were student-initiated or researcher-initiated. In each case, the process undergone and the result obtained were rated in terms of the success (in terms of achieving a correct response to the task) with which the question was answered and the efficiency of the process used by the participant in reaching the answer. In this context, efficiency was defined in terms of the number of operations undergone in order to reach the answer compared with the minimum number of operations required to reach the answer as determined by an experienced operator.

In this way, the success and efficiency with which the participants interacted with the databases was allocated a numerical value for each database, individually and for the aggregate of both databases.

5.6 Results

The statistical procedures appropriate to complete these analyses were limited by the number of cases in the study ($N=53$). As the data were arranged in nominal form, the most appropriate statistical procedure to use was the calculation of Spearman's correlation coefficients. This was achieved using the FREQ procedure in the SAS/STAT software package.

Analyses of relationships amongst the dependant variables were conducted at three levels:

1. for all students,
2. for year 8 students only, and
3. for year 12 students only.

The results for these analyses are summarised in tables 5.1, 5.2 and 5.3 (see appendix).

1. All students

Table 5.1 reveals, for all students in the study, a range of highly significant positive correlations. These are

- a) Success using the One World database is closely related to the students' efficiency using One World and total success and efficiency using both databases.

- b) Students' efficiency using One World is strongly related to their: success in using One World, success and efficiency using Volca, and total success and efficiency using both databases.
- c) Their success using Volca is strongly related to their: success using One World, success and efficiency using Volca, and total success and efficiency using both databases.
- d) The students' efficiency using Volca was strongly related to their: efficiency using One World, success using Volca, and total success rate using both databases revealed strong relationships with: success and efficiency using One World, success and efficiency using Volca and total efficiency on both databases.
- e) Finally, their efficiency using both databases was strongly related to their: success and efficiency using One World, success and efficiency using Volca, and success using both databases.

2. Year 8 students

The results in table 5.2 reveal that for year 8 students there is a similar pattern of relationships to the overall pattern revealed in table 5.1. Again there was no significant relationship between success using One World and success and efficiency using Volca. However, there was a significant relationship between success on One World and efficiency using Volca for the year 8 students.

3. Year 12 students

The results for the year 12 students, as revealed in table 5.3, do not provide as strong a set of positive relationships. In fact, a number of low, insignificant negative relationships were revealed between success and efficiency with both databases. Significant positive relationships between the dependent variables was found for the following pairings of data:

- (a) success using One World with efficiency using the same database;
- (b) efficiency using One World with efficiency using both databases;
- (c) success using Volca with overall success using both databases;
- (d) efficiency using Volca with overall efficiency using both databases; and
- (e) overall efficiency on the databases both separately and together.

The second set of analyses were the calculations of Spearman's correlation coefficients between each of the dependent variables and each of the ten independent variables. The results and significance of these calculations are summarised in table 5.4.

Significant positive relationships were found between the following variables:

- 1. chronological age of all students with success and efficiency using Volca and with overall efficiency in using both databases;
- 2. use of computers at home by year 12 students with success using Volca;
- 3. verbal intelligence for all students with success using One World; and
- 4. verbal intelligence for year 8 students with success and efficiency using One World, efficiency using Volca, and overall success and efficiency using both databases.

Significant negative relationships were found only for year 12 students. These relationships occurred for:

1. the correlations of general attitudes to computers with success in using Volca, and
2. previous use of databases with efficiency in using Volca and overall efficiency in using both databases.

5.7 Discussion and conclusion

The results of this study indicate a number of significant relationships between adolescents' success and efficiency in using menu-driven and command-driven geographical databases. While these results do not explain the nature of the relationships, their statistical significance does suggest that they are worthy of closer investigation. They do suggest that differences do occur when students use menu driven and command driven database.

Overall, the students in the study group were both more successful and more efficient using a menu-driven database. While this could have been due to some students having a general competence with databases derived from previous use, there is no evidence from the results to support this view. Year 8 students demonstrated a more consistent relationship between success and efficiency when using either menu driven or command driven databases in geography than did the year 12 students. Explanations for the performances of both groups of students may relate to the nature of each sample. They may relate to the range of experiences that both groups of students brought to the learning situation. The year 8 students tended to be quite active in exploring the capacities of the computers through trial and error. The year 12 students, through their additional years of formal education, were less willing to think laterally or to risk failure and so they had greater difficulty in handling the different types of database successfully and efficiently.

On both databases, success in performing the prescribed tasks presupposed efficiency in using such databases. This result supports the conclusion by Brazier & Trimp (1988) that well structured mental models are necessary if students are to use databases well. While Brazier & Trimp showed that if students were taught the mental model of the database they would find command driven databases better to use, we are suggesting that many of today's geography classes are not taught such models. In the absence of such models, menu driven databases assist the students to acquire appropriate mental models which allow them to use geographical databases both successfully and efficiently.

Several questions raised by this research remain to be investigated in future studies. Although we have considered students' prior access to computers both in school and at home, it is unclear whether total time spent practising with computers and software is influential on students' success and efficiency using geographical databases. The data reveals a general lack of correlation between spatial intelligence and almost all the efficiency and success measures in this study. Further studies may be needed to confirm that this is in fact the case or whether spatial intelligence is being swamped by the influence of verbal intelligence. Perhaps

qualitative observational studies of a similar group of students engaged in similar tasks would provide richer data on students' use of curriculum specific databases.

References

- BLISS, J., P. CHANDRA & M. COX (1986), The introduction of computers into a school. *Computers & Education* 10, pp. 49-54.
- BLOOM, B. (1956), *Taxonomy of educational objectives: Handbook 1; Cognitive domain*. New York: Mackay.
- BRAZIER, F. & H. TRIMP (1988), Teaching information retrieval skills: How instructional metaphors influence the effectiveness of interaction. Lovis & Tagg, eds., *Computers in Education*. Amsterdam: Elsevier Science Publishers B.V.
- DEGL'INNOCENTI, R. & M. FERRARIS (1988), Database as a tool for promoting research activities in the classroom: an example in teaching humanities. *Computing Education* 12 (1), pp. 157-162.
- DELCOMS, V. & S. KULEWICZ (1986), Improving computer based problem solving training: the role of the teacher as mediator. *Computers in Human Behaviour* 2, pp. 135-146.
- HURD, S. (1987), Computer use in economics. In: W. Kent & R. Lewis, eds., *Computer assisted learning in the humanities and social sciences*. Oxford: Blackwell Scientific.
- JOHNSON, V. (1987), The evaluation of microcomputer programs: an area of debate. *Journal of Computer Assisted Learning* 3 (1).
- LIDSTONE, J. (1989), The use of computers in geography teaching (Ch. 21). In: J. Fien et al., (1989) *The geography teachers guide to the classroom*. Second edition. Melbourne: Macmillan.
- LIDSTONE, J. & R. GERBER (1989), Australian courseware in geographical education (Paper presented to the IGU conference on geographical courseware). Freiburg (ERIC Clearinghouse for Social Studies/Social Science Education: ED 310056).
- MORAN, T. (1981), The command language grammar: a representation for the user interface of interactive computer systems. *International Journal of Man-Machine Studies* 15, pp. 3-50.
- NORMAN, D. (1983), Some observations on mental models. In: D. Genter & A. Stevens, eds., *Mental models*. Hillsdale, N.J.: Lawrence Erlbaum.
- SOMMERVILLE, I. (1982), *Software engineering*. London: Addison Wesley.
- UNDERWOOD, J. (1988), An investigation of teacher intents and classroom outcomes in the use of information handling packages. *Computer Education* 12 (1), pp. 91-100.
- WATSON, D. (1990), The classroom vs the computer room. *Computers & Education* 15 (1 3), pp. 33-37.

Appendix

Table 5.1 *Correlation analyses on success and efficiency ratings for selected geography databases using Spearman's correlation co-efficient*

(All Students)

	WC.SUC	WC.EFF	D.SUC	D.EFF	TOT.SUC	TOT.EFF
WC.SUC	1.0000 (0.0)	0.5672 (0.0001)	0.0586 (0.7122)	0.1970 (0.2111)	0.5663 (0.0001)	0.5334 (0.0003)
WC.EFF	*	1.0000 (0.0)	0.5443 (0.0002)	0.5444 (0.0002)	0.7363 (0.0001)	0.9328 (0.0001)
D.SUC	*	*	1.0000 (0.0)	0.7554 (0.0001)	0.7576 (0.0001)	0.6113 (0.0001)
D.EFF	*	*	*	1.0000 (0.0)	0.7091 (0.0001)	0.7586 (0.0001)
TOT.SUC	*	*	*	*	1.0000 (0.0)	0.8156 (0.0001)
TOT.EFF	*	*	*	*	*	1.0000 (0.0)

WC.SUC	= Success on World Countries database
WC.EFF	= Efficiency using World Countries database
D.SUC	= Success on Volcanoes database
D.EFF	= Efficiency using Volcanoes database
TOT.SUC	= Success using both databases
TOT.EF	= Efficiency using both databases

Table 5.2 *Correlation analyses on success and efficiency ratings for selected geography databases using Spearman's correlation co-efficient*

(Year 8)

	WC.SUC	WC.EFF	D.SUC	D.EFF	TOT.SUC	TOT.EFF
WC.SUC	1.0000 (0.0)	0.7088 (0.002)	0.2673 (0.2174)	0.5254 (0.0100)	0.7899 (0.0001)	0.7490 (0.0001)
WC.EFF	*	1.0000 (0.0)	0.7051 (0.0002)	0.5948 (0.0028)	0.8520 (0.0001)	0.9751 (0.0001)
D.SUC	*	*	1.0000 (0.0)	0.7943 (0.0001)	0.7572 (0.0001)	0.7676 (0.0001)
D.EFF	*	*	*	1.0000 (0.0)	0.8498 (0.0001)	0.7116 (0.0001)
TOT.SUC	*	*	*	*	1.0000 (0.0)	0.9389 (0.0001)
WC.EFF	*	1.0000 (0.0)	0.7051 (0.0002)	0.5948 (0.0028)	0.8520 (0.0001)	0.9751 (0.0001)
D.SUC	*	*	1.0000 (0.0)	0.7943 (0.0001)	0.7572 (0.0001)	0.7676 (0.0001)
D.EFF	*	*	*	1.000 (0.0)	0.8498 (0.0001)	0.7116 (0.0001)
TOT.SUC	*	*	*	*	1.0000 (0.0)	0.9389 (0.0001)
TOT.EFF	*	*	*	*	*	1.0000 (0.0)

WC.SUC = Success on World Countries database
 WC.EFF = Efficiency using World Countries database
 D.SUC = Success on Volcanoes database
 D.EFF = Efficiency using Volcanoes database
 TOT.SUC = Success using both databases
 TOT.EFF = Efficiency using both databases

Table 5.3 Correlation analyses on success and efficiency ratings for selected geography databases using Spearman's correlation co-efficient

(Year 12)

	WC.SUC	WC.EFF	D.SUC	D.EFF	TOT.SUC	TOT.EFF
WC.SUC	1.0000 (0.0)	0.7795 (0.0001)	-0.2587 (0.2848)	-0.1810 (0.4581)	0.3174 (0.1854)	0.3780 (0.1105)
WC.EFF	*	1.0000 (0.0)	-0.2228 (0.3592)	-0.0369 (0.8805)	0.1821 (0.4555)	0.5928 (0.0075)
D.SUC	*	*	1.0000 (0.0)	0.3298 (0.1678)	0.6313 (0.0037)	0.1061 (0.6654)
D.EFF	*	*	*	1.0000 (0.0)	0.2964 (0.2178)	0.7586 (0.0002)
TOT.SUC	*	*	*	*	1.0000 (0.0)	0.4503 (0.0530)
TOT.EFF	*	*	*	*	*	1.0000 (0.0)

WC.SUC = Success on World Countries database

WC.EFF = Efficiency using World Countries database

D.SUC = Success on Volcanoes database

D.EFF = Efficiency using Volcanoes database

TOT.SUC = Success using both databases

TOT.EFF = Efficiency using both databases

Table 5.4 Correlation analyses between success and efficiency ratings for selected geography databases and selected independent variables using Spearman's correlation co-efficient

	WC.SUC	WC.EFF	D.SUC	D.EFF	TOT.SUC	TOT.EFF
AGE						
Yr. 8	-0.2059 (0.3458)	0.0072 (0.9739)	0.2496 (0.2507)	0.1254 (0.5684)	-0.0254 (0.9081)	0.0035 (0.9873)
Yr.12	0.2105 (0.2105)	-0.0445 (0.8562)	0.4569 (0.0492)	-0.1739 (0.4670)	0.2150 (0.3766)	-0.2256 (0.3529)
All	-0.0587 (0.7115)	0.3882 (0.0111)	0.4459 (0.0031)	0.4670 (0.0018)	0.2734 (0.0798)	0.4209 (0.0055)
COMPEXP						
Yr.8	0.2684 (0.2156)	0.1000 (0.6490)	-0.0540 (0.8066)	0.0933 (0.6718)	0.0579 (0.7927)	0.0613 (0.7810)
Yr. 12	0.3617 (0.1401)	0.2958 (0.2333)	-0.4658 (0.0514)	-0.1352 (0.5926)	-0.2311 (0.3560)	0.0916 (0.7177)
All	-0.2942 (0.0619)	0.3026 (0.0545)	-0.1057 (0.5106)	0.1045 (0.5156)	0.0824 (0.6084)	0.2607 (0.0996)
COMPATT						
Yr. 8	0.0573 (0.7948)	-0.0326 (0.8824)	-0.2646 (0.2223)	-0.1698 (0.4384)	-0.2218 (0.3090)	-0.1455 (0.5076)
Yr. 12	0.2334 (0.3511)	0.2184 (0.3839)	-0.6533 (0.0033)	-0.2689 (0.2804)	-0.3862 (0.1134)	-0.0448 (0.8599)
All	0.0775 (0.6523)	0.3110 (0.0477)	-0.1057 (0.5104)	0.1141 (0.4772)	0.0081 (0.9595)	0.2908 (0.0651)
H.COMP						
Yr. 8	-0.1762 (0.4210)	-0.0638 (0.7722)	0.0166 (0.9400)	-0.0500 (0.8205)	-0.0480 (0.8276)	-0.0644 (0.7702)
Yr. 12	-0.4142 (0.0874)	-0.5063 (0.0320)	0.4642 (0.0523)	-0.0230 (0.9278)	0.2282 (0.3624)	-0.3651 (0.1362)
All	-0.2848 (0.0711)	-0.2845 (0.0714)	0.1080 (0.5014)	-0.0375 (0.8156)	-0.0978 (0.5426)	-0.2615 (0.0986)
H.USE						
Yr. 8	-0.0040 (0.9855)	-0.0019 (0.9928)	-0.0796 (0.7180)	-0.1666 (0.4472)	-0.0670 (0.7613)	-0.0301 (0.8913)
Yr. 12	-0.4760 (0.0458)	-0.3756 (0.1245)	0.7276 (0.0006)	0.0715 (0.7780)	0.2772 (0.2653)	-0.2582 (0.3008)
All	-0.1663 (0.2987)	-0.2819 (0.0714)	0.0199 (0.9017)	-0.1565 (0.3282)	-0.1354 (0.3986)	-0.2962 (0.0600)
SCH.EXP						
Yr. 8	0.2356 (0.2790)	-0.0002 (0.9921)	-0.3224 (0.1335)	-0.0607 (0.7831)	0.0189 (0.9317)	-0.0185 (0.9332)
Yr. 12	-0.0638 (0.8013)	0.2127 (0.3966)	0.3390 (0.1687)	0.0255 (0.9199)	0.2236 (0.3723)	0.0843 (0.7392)
All	0.0406 (0.8008)	0.1643 (0.3045)	0.0231 (0.8859)	0.1049 (0.5136)	0.1257 (0.4333)	0.1761 (0.2707)

Table 5.4 Correlation analyses between success and efficiency ratings for selected geography databases and selected independent variables using Spearman's correlation co-efficient (continued)

	WC.SUC	WC.EFF	D.SUC	D.EFF	TOT.SUC	TOT.EFF
SCH.USE						
Yr. 8	0.0776 (0.7248)	0.0768 (0.7275)	-0.0248 (0.9104)	0.0701 (0.7504)	0.0925 (0.6473)	0.0893 (0.6853)
Yr. 12	0.0230 (0.9277)	0.0534 (0.8333)	-0.0147 (0.9536)	-0.2303 (0.3577)	0.1703 (0.4993)	-0.0913 (0.7184)
All	0.0615 (0.7025)	-0.0920 (0.5672)	-0.1522 (0.3419)	-0.2194 (0.1681)	-0.0865 (0.5906)	-0.1293 (0.4204)
DBASE						
Yr. 8	0.0247 (0.9107)	0.2365 (0.2771)	0.1294 (0.5562)	0.1240 (0.5728)	0.1160 (0.5981)	0.1793 (0.4130)
Yr. 12	0.1519 (0.5500)	-0.2218 (0.3762)	-0.2840 (0.2534)	-0.6532 (0.0033)	-0.1965 (0.4343)	-0.6411 (0.0041)
All	0.0372 (0.8170)	0.2182 (0.1703)	0.0563 (0.7263)	-0.0420 (0.7940)	0.1242 (0.4388)	0.1167 (0.4673)
SP.INT						
Yr. 8	0.4319 (0.0447)	0.0958 (0.6712)	-0.2403 (0.2813)	0.1006 (0.6559)	0.1739 (0.4387)	0.1130 (0.6166)
Yr. 12	0.1268 (0.6275)	0.2933 (0.2532)	-0.1118 (0.6692)	-0.2954 (0.2495)	0.0242 (0.9265)	-0.0528 (0.8405)
All	0.3231 (0.0448)	0.2474 (0.1288)	-0.0966 (0.5583)	0.0069 (0.9669)	0.1235 (0.4536)	0.1832 (0.2642)
V.INT						
Yr. 8	0.7905 (0.0001)	0.5866 (0.0041)	0.2884 (0.1929)	0.5417 (0.0092)	0.7037 (0.0003)	0.6599 (0.0008)
Yr. 12	-0.0269 (0.9273)	0.1984 (0.4965)	-0.1912 (0.5125)	0.0033 (0.9909)	-0.2371 (0.4142)	0.0100 (0.9729)
All	0.4840 (0.0028)	0.2048 (0.2308)	0.0660 (0.7017)	0.1509 (0.3795)	0.2799 (0.0982)	0.1845 (0.2812)

- WC.SUC = Success on World Countries database
- WC.EFF = Efficiency using World Countries database
- D.SUC = Success on Volcanoes database
- D.EFF = Efficiency using Volcanoes database
- TOT.SUC = Success using both databases
- TOT.EFF = Efficiency using both databases
- AGE = Chronological age
- COMPEXP = Perceived general experience with computers
- COMPATT = General attitude to computers
- H.COMP = Availability of computer in the home
- H.USE = Use of computer in the home
- SCH.EXP = School subjects in which students used computers
- SCH.USE = Extent of computer use at school
- DBASE = Previous use of databases
- SP.INT = Spatial Intelligence
- V.INT = Verbal Intelligence

6 GEOGRAPHICAL PROCEDURAL KNOWLEDGE AND MAP SKILLS

Joop van der Schee, Henk van Dijk & Hans van Westrhenen

6.1 Introduction

In the last few years Dutch educational policy has been strongly emphasizing the importance of improving the quality and the efficiency of education. In connection with this it is important to find out what specific subject differences there are between students and how to best organize education to minimize those differences. In 1991 a new research program was started at the Free University at Amsterdam to investigate this issue. This research program is called "Curriculum studies and methodological differentiation in secondary education". The research is directed towards two main questions:

- A. What differences in subject knowledge and skills do secondary school students show when they perform subject specific tasks?
- B. Which methodological approaches help to improve the quality of education in the sense that more students achieve the goals formulated beforehand? These are the guiding questions for the research being done for four school subjects: geography, biology, French and physics. This article is a report on the research design and on the first research results for geography.

6.2 Procedural knowledge

The general objective of Dutch geography teaching for students between 12 and 15 years of age is: "Geography teaching is designed to equip young people with skills which will enable them to participate in various social and life situations in the spatial relationships of which they are a part: their own environment, the Dutch, the European and the world community" (Commissie Herziening Eindtermen 1990, p.54). In other words, students need to acquire skills that will help them perform in a wide variety of regional frameworks. With regard to this, there is a pertinent conclusion in Van Westrhenen's dissertation (1976): "As a result of the descriptive character of the learning material for geography, students are offered only a very limited set of instruments to help them, from a geographical point of view, make sense of the many phenomena they are confronted with in daily life". The research presented here investigates the differences that students manifest in answering questions for which such a set of geographic instruments is needed and the manner by which this set of geographic instruments can be taught successively. The set of instruments consists of geographical procedural knowledge. We will briefly discuss

here what is meant by declarative and procedural knowledge in educational geography.

In geography declarative knowledge consists of facts and generalizations. A geographical statement is of a factual nature when it concentrates on the spatial object as such. For instance, "The Netherlands have 15 million inhabitants". If a geographical statement concerns one or more classes of spatial objects, then we are talking about geographical generalizations. Examples of these are "The countries in the Middle East are mainly Islamic" and "In the various regions where the amount of precipitation is higher in Morocco, the population density is greater". Connected generalizations form comprehensive knowledge (Van Westrhenen 1989a). Procedural knowledge in geography concerns either the actions which are necessary to get from geographical facts to geographical generalizations, or those necessary to explain or predict facts using the geographical generalizations offered. In other words, both inductive and deductive approaches are possible. In our research subject skills are defined as geographical procedural knowledge. Maps play an important role in this. The rest of this paragraph will look at this more closely.

Geography teaching presents knowledge in various ways: texts, tables, graphs and maps. For the geographer the map is the most important of these various sources of information. The added value of a map is that the location of all of the phenomena and areas is known. Maps show lots of data and in this way offer the possibility of both finding locations, spatial distributions and areal differentiations and the areal associations between these spatial distributions and areal differentiations, which are important in order to achieve more insight into reality. What a user sees on a map and can conclude from that depends on his ability to use map skills. Map skills can be divided into three stages of action:

1. Map reading. This includes identifying and naming phenomena on a map.
2. Map analysis. This includes the recognition of spatial distributions, areal differentiation and areal associations on a map.
3. Map interpretation. This includes making statements or predictions using the areal associations discovered on the map. To do this one needs other sources of information than just the map.

In all three stages of action there is procedural knowledge. Someone who cannot use the legend and scale, cannot identify phenomena on a map (map reading). Someone who cannot classify facts into spatial distributions and areal differentiations and cannot discover areal associations on a map is not able to analyze a map. The map user will only be able to interpret the map if he has enough procedural and declarative knowledge.

Someone with limited map skills will generally only be able to describe the individual objects on the map. For example looking at the map in figure 6.1, he could say "Grenoble has the largest population", or "Aix-les-Bains is lower than 800 meters". There is an endless number of similar geographical facts that could be gleaned from the map. Greater map skills are necessary to organize information on the map, which makes it possible to discover that relationships can be discovered between various characteristics of areas. The following remarks can be made concerning the spatial distribution in the map in figure 6.1 (Van der Schree 1987, p.69):

- a. Areas in the western part of the Alpes du Nord have, on average, more inhabitants than areas in the eastern part.
- b. Areas in the western part of the Alpes du Nord are on average lower than areas in the eastern part.

In both statements collections of geographical phenomena (in this case, areas) are compared as to certain site and situation characteristics. In order to make the statements mentioned above, the map user places an imaginary grid pattern over the map (Spruit 1974, p. 11). For every region on the map he studies both the situation, the population density and the altitude. In this way the area on the map can be spatially segmented. One can conclude that many areas in the west of the Alpes du Nord have many inhabitants and are situated at a low altitude, while for many areas in the west the opposite is true. Using these statements we can come to a generalization. Areas lower in altitude have more inhabitants than areas higher in altitude.

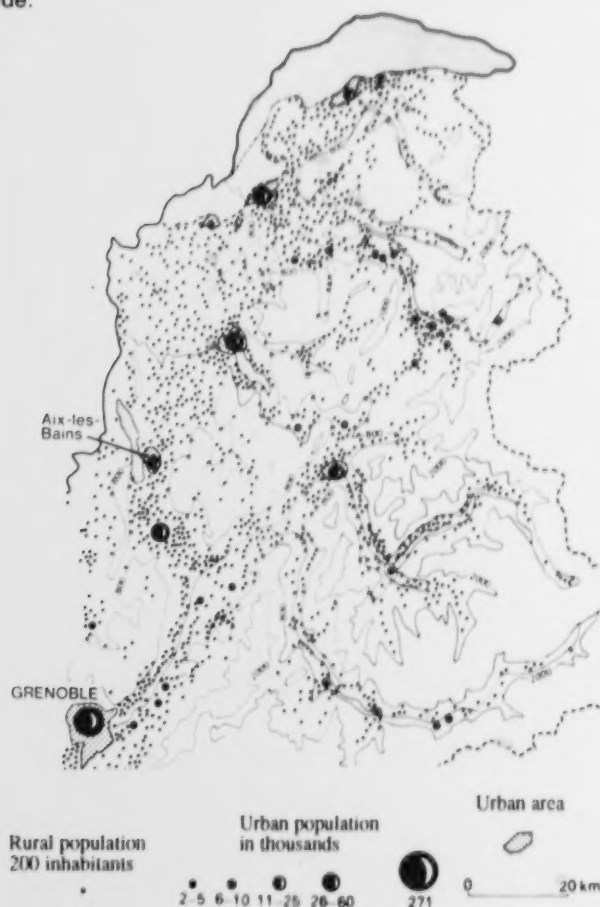


Figure 6.1 Population distribution in Alpes du Nord (Beaujeu-Garnier 1976)

It must be taken into account here that the size of the areas chosen are of great influence in determining the content and the strength of the generalizations. Geographers attempt to find general patterns. According to Hoekveld (1989): "The geographer looks for general and specific characteristics of areas. First one looks at the whole and that which can be generalized and subsequently one turns to the particular".

6.3 Research questions

As mentioned at the beginning of this article, this research is connected to a larger research program. The two main research questions for geography are:

1. What are the differences between students in the application of geographical procedural knowledge when they perform subject specific tasks?
2. How should we structure and differentiate geographical lesson materials based on geographical procedural knowledge - deductively or inductively - in order to achieve improvement in the execution of subject specific tasks for more students as far as achieving aims set beforehand is concerned?

In this research project learning is seen as a way of information processing. Factual information is processed into comprehensive knowledge which in turn is used to understand new facts or to make predictions. As far as learning processes are concerned this research is connected to the cognitive approach to learning processes as it is described by, among others, Boekaerts (1987) and Boonman & Kok (1982). Knoers (1973) states that education will prove more efficient when in addition to factual knowledge there is also comprehensive knowledge. Comprehensive knowledge offers possibilities for more and different kinds of achievement, because this kind of knowledge is more flexible than factual knowledge (Van Parreren 1974). The research presented here continues the educational geographical research that has taken place at the Free University since the 1970's (Van Westrhenen 1989b). A theoretical basis for this research can be found in Posner & Strike's model (1976). This model assumes that the 'inquiry process' is an indispensable link between the empirical phenomena and the conceptual systems that students have to acquire (Van Westrhenen 1989b). Possessing geographic procedural knowledge (the 'inquiry process') seems to be an important factor in structuring and sequencing differentiated learning in secondary education (Hoekveld-Meijer & Schutte 1988; Van der Schee 1987; Van Westrhenen 1989a).

6.4 The research method

This research has been divided into two stages, an explorative stage and a testing stage. This is true for both research questions. The explorative stage will be used to find differences between students in subject skills and also to gain experience with the effects on learning achievements of lesson materials that want to train students explicitly in geographical skills that are needed to perform subject specific tasks. Using testing research we will attempt to make significant conclusions about differences between students in subject skills and about the effects of systematic training in subject skills. It will only be possible to formulate hypotheses after the exploration stage has been completed.

The research for the first research question has the following independent variables: age, sex, general spatial insight, reasoning ability, educational motivation, motivation concerning geographical topics and regions and school achievement (the most recent report grades for geography, languages, maths, physics and chemistry). These variables are measured by the scores of the students on the so-called student-variable test. The dependent variables in the first research question are the skills students have in identifying, classifying and relating spatial distributions and areal differentiations. These skills are measured by the students' scores on a geographical test which will later serve as a pretest for the second research question. The research concerning the second research question contains the following independent variables (also see figure 6.2): the students' scores on the variables mentioned above in the student-variable test and the scores of the students on the pre-test.

Condition					
Group 1:	LVT	VT	X	NT	RT
Group 2:	LVT	VT	Y	NT	RT
Group 3:	LVT	VT	Z	NT	RT

Explanation of the abbreviations:

LVT	=	student variable test
VT	=	pre-test
X	=	inductive learning mode
Y	=	deductive learning mode
Z	=	control condition
NT	=	post-test
RT	=	retention test

Figure 6.2 Research design for research question 2

The conditions for research question 2 are:

- Experimental condition X: geographical lesson materials with regularly occurring exercises in subject skills with an emphasis on map skills. Modules which are exclusively aimed at training map skills are alternated with modules in which subject skills are included in texts and assignments concerning certain areas (e.g. Europe) and topics (e.g. tourism). The learning mode is primarily inductive;
- Experimental condition Y: the same as type X, with the exception that the learning mode is primarily deductive;
- Control condition Z: geographical lesson materials about the same subject which is taken from a current geography textbook.

The dependent variables for research question 2 are the student scores on the post-test and the student scores on the retention test. The total number of class hours that goes with the design of research question 2 is about 15, including the administering of the tests.

The research design for research question 2 can be described as a "pretest-posttest design with treated control group" (Cook & Campbell, 1979). Random treatment is difficult in this research. So this might be termed a "quasi-experiment". Any progression between the pretest and the posttest will primarily be measured in terms of more students who have achieved the pre-determined aims. In addition, the student scores on the pretest and the posttest will be compared and time needed for learning will be taken into account.

Learning aims will be based on the current graded objectives for geography at school. The pretest, the posttest and the retention test will, in order to meet the requirement of comparability, test the same geographical procedural knowledge, although the declarative knowledge offered will differ. During the explorative beginning stage of the research thinking-aloud protocols will be used. For the student variables 'general spatial insight', 'reasoning ability' and 'schoolmotivation' use will be made of existing tests. A number of observations concerning the validity and reliability will be made here.

In the explorative stage the validity of the analysis will be ensured by using a variety of research instruments. Reliability increases if the analyses of the thinking aloud protocols show the same tendencies as the scores on the various written tests. The content validity in this stage will be guaranteed by discussing the research instruments with various groups of teachers and experts. A stable response group of geography teachers will be asked for their remarks and observations at regional meetings. In addition, a randomly chosen group of geography teachers will be asked to provide a written evaluation. The written tests are very important in this project, especially in the testing part of the research. As far as we know, validated and reliable geography tests designed to measure skills in the fashion described here are nonexistent. We will attempt to make uniform subtests with a high reliability coefficient. In principle it is possible to design separate subtests for the various geographical skills. This kind of research may be typified as class-directed research, according to Boekaert's (1983, p.68) classification. Since this research concerns empirical research in a natural educational setting, so-called 'ecologically valid' research (De Klerk 1983), there is bound to be a certain amount of tension between external and internal validity.

6.5 The research population and time planning

We intend to answer the research questions by doing our research with 7th to 10th graders. The research will be directed at classes with a heterogeneous classroom population (12 to 15 years of age). This with a view to introducing differentiation into subject specific tasks. Children just starting at secondary school will be especially looked at. On a limited scale students at the top end of the age range (grade 10) will be considered in the research in order to be able to make statements concerning developments that are age-related and (subject) experience-related. In the exploratory stage of the research data will be collected about individual students or small groups of students. The testing part of the research will take place in 12 heteroge-

neous classes in grades 7 to 10 at secondary schools. More than 300 pupils will be involved.

The influence of the teachers who will be teaching the geography lessons will be reduced by working with teachers who have a number of parallel classes.

It will take five years to do the entire research project. In the first stage map skills tests will be developed, which will be used to research with a view to answering the first research question.

At first the research will be mainly qualitative, with an aim to developing a good measuring instrument. Subsequently, the research with this measuring instrument will become more quantitative as it is used to discover significant differences between students.

In the second research stage a number of lesson modules will be developed which will be used for the second research question. Following the exploratory research with the lesson materials, the research design as shown in figure 6.2 will be performed in 1994. The final report is planned for 1995.

6.6 Exploratory research

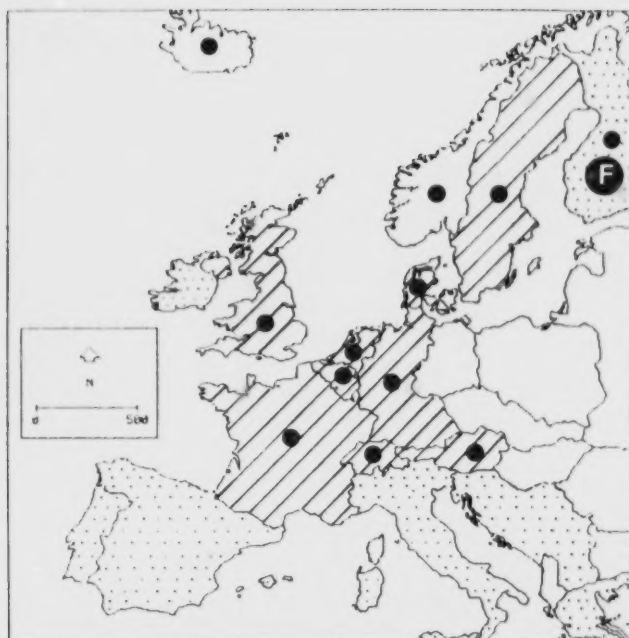
A. The design of the first exploratory research.

In the spring of 1991 a first version of a map skills test was designed. The main aim of this test is to find out whether students think mostly in an object-oriented fashion or whether they are also able to recognize relationships between spatial distributions and areal differentiations on a map. The test consists of five questions which each require a different geographical skill.

For practical reasons we will look more closely here at the first two questions of the map skills test, so the first two questions of the test are presented here more extensively than the other three.

Figure 6.3 shows the first question, in which students are asked to say why, according to the information on the map, Finland is an unusual country. In order to be able to give the right answer to this question the students need to look at the entire map and find associations between the two areal differentiations on the map. Figure 6.4 shows a question concerning spatial distribution. For this question about France students have to indicate the best location to build a new nuclear reactor and to explain why they choose A, B, C or D. The second part of this assignment concerns the location for a new electrohydraulic dam. In order to be able to answer this question properly students need to discover that all of the nuclear reactors are located close to water (connecting two spatial distributions) and that all of the hydroelectric dams are located in rivers and also in places with large drops in elevation (connecting a spatial distribution to an other spatial distribution and an areal differentiation).

Figure 6.5 shows the third question, in which students are asked to say at which location, according to the information on the maps, a rice growing experiment will succeed. The three maps and the assignment of the fourth question are shown in figure 6.6. For this question students have to decide what the best plan will be to build three new large hypermarkets. Figure 6.7 shows the fifth question, in which students are asked to say why, according to the information on the map, it is strange that country Y had many tourists from country X and country Z had not.



- rich country (country where the average income per capita was more than 20.000 guilders in 1983)
- ▤ country from which at least 100.000 workers left to work in another country in 1983
- ▨ country with at least 100.000 foreign workers in 1983
- no data available
- ⓕ FINLAND

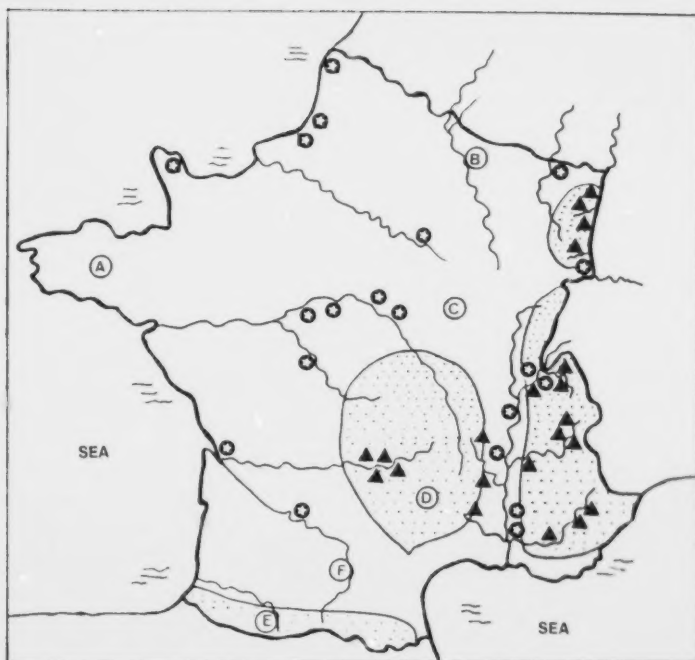
Question 1: According to the information on the map Finland is a country from which a lot of workers leave to work in another country and Finland is also a rich country. When you look at the map why is that unusual?

.....

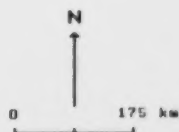
.....

.....

Figure 6.3 Question 1 about Finland from the map skills test (inductive open version)



- ★ nuclear reactor
- ▲ hydroelectric dams
- ▨ area higher than 500 m. with large drops in elevation
- ~~~~~ river



Question 2A: The French government is planning to build a new nuclear reactor. When you look at the map what do you think the best location is to build this nuclear reactor?

A, B, C, or D? because:

Question 2B: The French government is also planning to build a new hydroelectric dam. When you look at the map what do you think the best location is to build this hydroelectric dam?

A, D, E or F? because:

Figure 6.4 Question 2 about France from the map skills test (inductive open version)



BERGEZIË

HEIGHT



MOUNTAINS (higher than 500 meters)



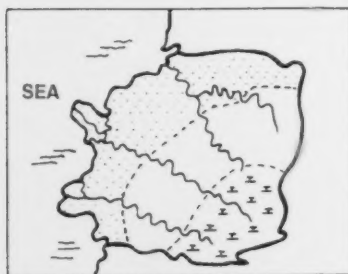
HILLS (200-500 meters)



LOWLAND (lower than 200 meters)



RIVER



BERGEZIË

CLIMATE



COLD AND SOMETIMES MOIST



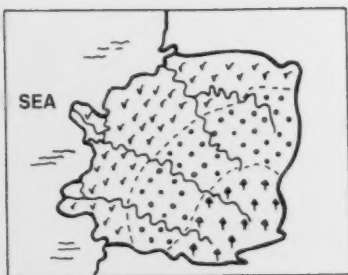
HOT AND DRY



HOT AND MOIST



RIVER



BERGEZIË

LAND USE



PINE-FOREST



GOATS AND SHEEP



RICE

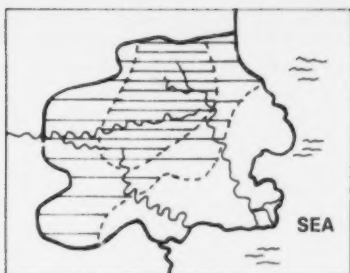


RIVER



Figure 6.5

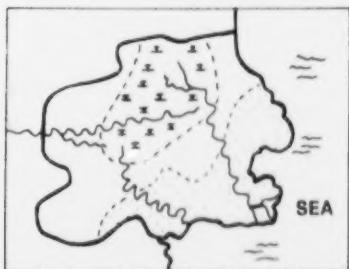
Question 3 about Bergezië and Geria from the map skills test
(inductive open version)



GERIA

HEIGHT

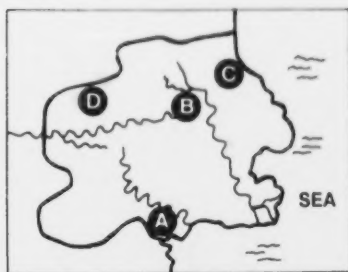
-  MOUNTAINS (higher than 500 meters)
-  HILLS (200-500 meters)
-  LOWLAND (lower than 200 meters)
-  RIVER



GERIA

CLIMATE

-  COLD AND SOMETIMES MOIST
-  HOT AND DRY
-  HOT AND MOIST
-  RIVER



GERIA

LAND USE

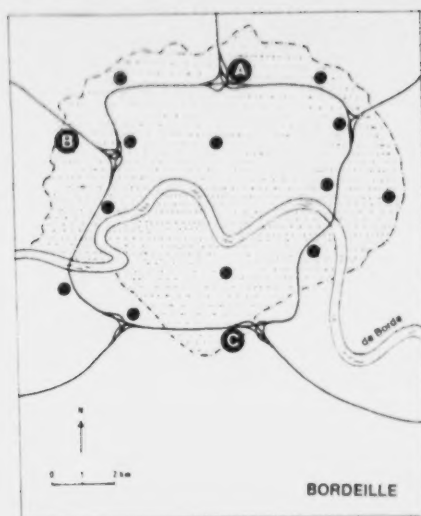
-  RIVER



Question 3: Look carefully at the three maps of Bergezië. Then, look at the three maps of Geria. There are four letters on the last map of Geria. They indicate the locations where rice growing experiments have been made on farms. When you look at the maps, what do you think is the location where the experiment succeeded?

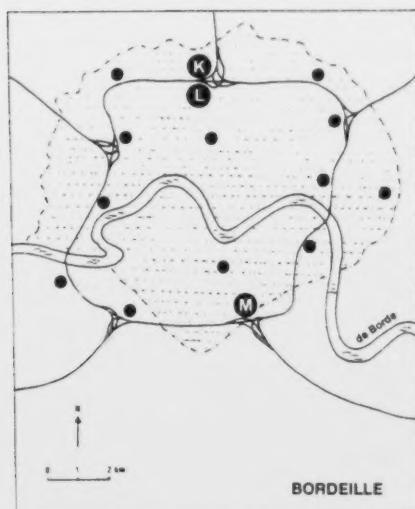
A, B, C or D? because:

Figure 6.5 Question 3 about Bergezië and Geria from the map skills test (inductive open version, continued)



MAP 1 : PLAN VAN DIJK

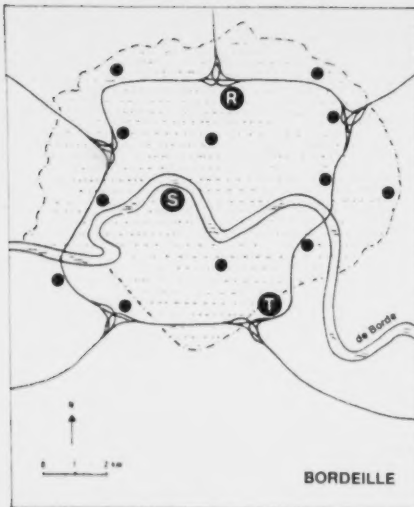
- PLANNED HYPERMARKET
- EXISTING HYPERMARKET
- ~ RIVER
- HIGHWAY
- - - CITY BORDER
- ▨ INHABITED AREA
(with an even distribution
of the population)



MAP 2 : PLAN JONES

- PLANNED HYPERMARKET
- EXISTING HYPERMARKET
- ~ RIVER
- HIGHWAY
- - - CITY BORDER
- ▨ INHABITED AREA
(with an even distribution
of the population)

Figure 6.6 Question 4 about Bordeaux from the map skills test (inductive open version)



MAP 3 : PLAN FARIAN

- PLANNED HYPERMARKET
- EXISTING HYPERMARKET
- ~ RIVER
- HIGHWAY
- - - CITY BORDER
- ▤ INHABITED AREA
(with an even distribution
of the population)

Question 4:

The three maps show the city of Bordeille. This city has 12 hypermarkets. Hypermarkets are very large supermarkets, sometimes as large as a couple of football grounds together. You can buy there food and drinks, but also clothes, books, household goods and so on. The pictures with this assignment show such a hypermarket. SUPER M plans to build 3 new very large hypermarkets in Bordeille. The three managers Van Dijk, Jones and Farian developed their own plan:

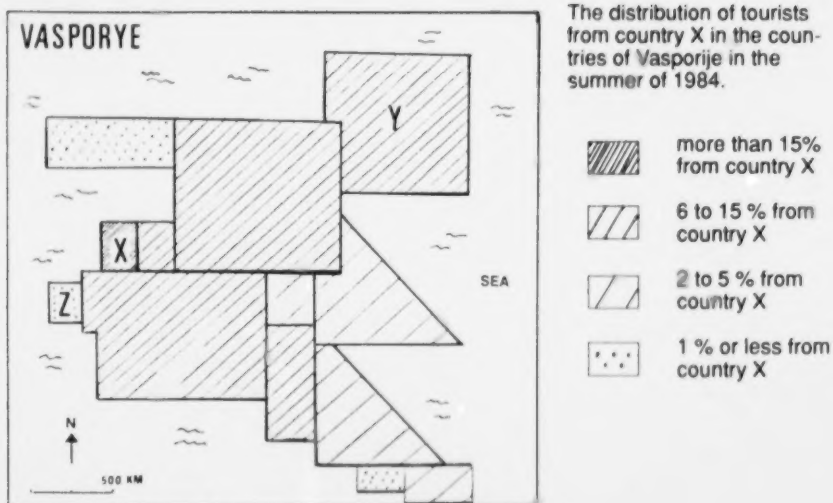
- Plan Van Dijk (see map 1): the three new hypermarkets have to be build at A, B and C.
- Plan Jones (see map 2): the three new hypermarkets have to be build at K, L and M.
- Plan Farian (see map 3): the three new hypermarkets have to be build at R, S and T.

When you look at the three maps what do you think is the best plan to build the new hypermarkets?

PLAN because



Figure 6.6 Question 4 about Bordeille from the map skills test
(inductive open version, continued)



Question 5:

The map shows the continent Vasporije. This continent consists of 13 countries. Each year a lot of people from country X go for a holiday abroad.

A: Many tourists from country X went to country Y in 1984; 10% of all tourists from country X went to country Y in 1984. When you look at the map why is that unusual?

.....

B: Not many tourists from country X went to country Z in 1984; 1% of all tourists from country X went to country Z in 1984. When you look at the map why is that unusual?

.....

Figure 6.7 Question 5 about Vasporije from the map skills test (inductive open version)

Earlier we claimed that geographical procedural knowledge is needed for both the inductive and the deductive learning modes.

For this reason we designed a variant for the map skills test in which the generalization in the map is not mentioned in the text and a variant in which it is in the text. The expectation is that the deductive version of the map skills test will score higher than the inductive one. In addition, a multiple choice version of both tests was designed to find out whether pre-structured answers would lead to better results. These are the four variants for the map skills test:

- A. Inductive version with open questions.
- B. Deductive version with open questions.
- C. Inductive version with multiple choice questions.
- D. Deductive version with multiple choice questions.

In June 1991, 108 students in four classes at two Dutch secondary schools participated in this first explorative research. These were all first year secondary school students (age 12 or 13).

One hundred students did the four variants of the map skills test in writing. Eight students were taken separately. They were given one of the tests by an interviewer and were asked to do the test while thinking aloud. Thinking aloud is a relatively young research method, and it is particularly effective in exploratory research concerning problem-solving (Breuker et al. 1986, p.241).

B. The results of the first explorative research

First a few brief remarks will be made concerning the results on the four written tests. After that we will look at the geographical content of the students' answers, including the thinking aloud protocols. Table 6.1 shows the scores of the students who took part in the map skills test. Those answers are counted correct which explicitly show that the student has used spatial distribution and areal differentiations on the map, as well as the associations between them, to solve the problem presented.

The assumption was made that the deductive version of the test would score higher than inductive one and that the closed version would score higher than the open version. Let us first look at the scores on the *inductive and deductive versions* of the map skills test in table 6.1. Variant D scores higher than C on all of the questions except question 5B. The supposition that the deductive version will score higher than the inductive version is confirmed for the closed version of the map skills test. However, when we compare the inductive open version and the deductive open version, we see that the deductive version does not score higher than the inductive version. If we take all of the questions for all 100 participating students together, the deductive version only has 2% more of all right answers than the inductive version. Our supposition is, therefore, not confirmed by this research.

An explanation for the relatively high score for the inductive open version as compared to the deductive open version may have something to do with the characteristics of the groups of students involved. The students that belong to the group that made the inductive open version are on average higher academic achievers than those of the other three groups.

Table 6.1 The percentage of right and wrong answers per question to the four variants of the map skills test

	Variant A inductive open (N=26)	Variant B deductive open (N=24)	Variant C inductive closed (N=26)	Variant D deductive closed (N=24)	Total (N=100)
<i>Question 1</i> (see fig. 6.3)					
right	23	4	38	46	28
wrong	77	96	62	54	72
<i>Question 2A</i> (see fig. 6.4)					
right	15	0	38	79	33
wrong	85	100	62	21	67
<i>Question 2B</i> (see fig. 6.4)					
right	12	0	31	63	26
wrong	88	100	69	37	74
<i>Question 3</i> (see fig. 6.5)					
right	65	54	69	75	66
wrong	35	46	31	25	34
<i>Question 4</i> (see fig. 6.6)					
right	42	50	42	75	52
wrong	58	50	58	25	48
<i>Question 5A</i> (see fig. 6.7)					
right	65	71	62	75	68
wrong	35	29	38	25	32
<i>Question 5B</i> (see fig. 6.7)					
right	62	63	81	75	70
wrong	38	37	19	25	30

When we compare the *open and closed versions* of the tests we see that the closed versions, as expected, showed an equal or higher percentage of right answers than the open versions (see table 6.2). When we take all of the questions for all of the 100 participants together we see that the closed versions contribute 24% more right answers than the open versions. It should be kept in mind that with the closed versions there is a possibility that students have guessed the right answer.

Looking now at the various questions in table 6.1, it appears that the first two questions score considerably lower than the other questions. One could ask whether questions one and two ask more geographical procedural knowledge of the students than the other questions. To a certain extent this is true. Questions 1 and 3 can both be classified as seeing associations between areal differentiations, but in question one there are a large number of smaller regions unlike question three. Questions 2 and 4 can be classified as questions concerning spatial distributions and areal differentiation. The complexity of the map in question 2, however,

is greater because the number of spatial distributions is larger than in question 4. However, this does not prove that the differences in subject content complexity are responsible for the different scores per question.

Table 6.2 The percentage of right answers in the open and closed versions of the map skills test as compared to the total number of right answers per question

	Open (N=50)	Closed (N=50)	Total (N=100)
Question 1	25	75	100
Question 2A	12	88	100
Question 2B	12	88	100
Question 3	45	55	100
Question 4	44	56	100
Question 5A	50	50	100
Question 5B	44	56	100
Average	38	62	100

In addition to the geographical content of the map information presented there is a second important factor of influence on the scores. Despite the fact that the students were explicitly asked to use the information on the map, the *effect of previous knowledge* on the part of the students is big. Depending on the problem presented, this can have various results. The effect of previous knowledge is very strong with the third question (figure 6.5). Many students do not analyze the maps, but just give answers on account of what they have learned before.

Table 6.1 shows that with *the first question* few students refer explicitly to the areal associations on the map when answering the question. On further analysis of the incorrect answers to the inductive open version and the deductive open version of question 1 it appeared that for 31% and 21% respectively of the students it is unclear whether they just looked at one object or considered the entire collection of phenomena. An example of such an unclear answer is: "If Finland is such a rich country, why do so many laborers go to another country to work?" In these cases it seems that students answer the question based on their previous knowledge of the subject in question. There is no evidence at all from their answers that they have studied the map as a whole in order to give their answer.

In addition to these unclear answers there are, for the open version of the first question, two main types of incorrect and incomplete answers that need to be distinguished. On the one hand there are students who only repeat the information presented, on the other hand there are students who mention facts which do not give the desired answer to the question. An example of repeating a presented fact is: "Because Finland is rich". An example of mentioning correct but irrelevant facts is: "It is very cold there".

The qualitative research offers useful extra information about the written test. The protocols for our qualitative research show clearly that the reactions of many students to the issue they were presented with are dominated by associations that they make concerning a certain area or topic. A clear example of this is the reaction of Ralf, who, thinking aloud about question 1 in the deductive open version of the map skills test (see figure 6.3, too), said:

"Finland is actually an industrious area, very cold so you wouldn't expect many people to live there. It says here that a lot of people have left for abroad to work. I think that there is a lot of work in the country itself, too, in the lumber industry or something.....they can make furniture from that in factories".

Ralf associates Finland with many images none of which can be found on the map: cold, thinly populated, lots of jobs, a lumber industry and a furniture industry. Ralf has no shortage of declarative knowledge, and he seems to be able to use the generalization that cold areas are generally thinly populated. He also seems to know that forestry and the wood industry connected to it are important means of livelihood in Finland. Ralf was able to give this information without consulting the map.

Although the question explicitly leads in that direction, Ralf did not use the map to compare the features of countries. Comparing areas and their features - using geographical procedural knowledge - by using the map in front of him does not seem to be an obvious course of action to Ralf. The thinking aloud protocols show clearly that Ralf is not an exception. A number of students use previously learned declarative knowledge to find an answer to "Why is Finland a special country when you look at the map?". Most students do not use a systematic classification of the information provided in order to find out whether there is a connection between average income and foreign labor in the countries of Europe (as shown by table 6.3).

Table 6.3 *The relationship between foreign labor and average income on the map of the first question in the map skills test*

Country	Countries the foreign laborers come from	Countries the foreign laborers went to	Countries with no information about foreign laborers	Total
Rich	1	9	2	12
Poor	6	0	8	14
Total	7	9	10	26

Instead of determining the special position of Finland using the information on the map, lots of students mention features of Finland that they find special, such as:

- * *Finland is cold.*
- * *Finland has a wood industry.*
- * *Finland is a northern country.*
- * *Finland has few tourists.*
- * *Finland is a small country.*
- * *Finland has a language they are not familiar with.*
- * *Finland is underpopulated.*

From the information that the students give it often appears that in solving a problem they only use the map in the last instance. It is more usual to use previous knowledge or knowledge presented in the text. One student, for example, when the interviewer asks him how he knows that Finland is a rich country, answers "It's in the question, I read that". Like with the first question, the percentage of right answers for inductive and deductive open versions of *the second question* in the map skills test was very low (see table 6.1). Of the 50 students who did an open version of the map skills test, 44% gave answer "B is on a river" to question 2A without pointing out in their explanation that all nuclear reactors are located next to water. For question 2B 38% of the students answered "E is on a river in the mountains" without in their explanation using as a reason the fact shown on the map that all hydroelectric dams are located next to water in areas with a big drop in elevation. In addition to this category of answers there is another kind of answer of interest in light of the research being described here. This concerns answers to question 2A, answers like "They don't have a nuclear reactor in that area yet". In this case the student does use procedural knowledge, because as we can see the spatial distribution of nuclear reactors is taken into consideration. The student does not, however, connect the spatial distribution of nuclear reactors with the spatial distribution of water on the map, but thinks that nuclear reactors should not be located too closely to each other. The student is capable of recognizing spatial distribution, but not capable of recognizing areal associations. What we have here is incomplete procedural knowledge. One fifth of the students scores in this category for the open version of question 2A in the map skills test. Even fewer students - only 8% - appear to be able to explicitly mention the areal associationship on the map (see table 6.4). Because of the small number of students in this explorative research, it is difficult to make statements about the use of geographical procedural knowledge of the same students with different questions on the test.

Table 6.4 Percentage distribution of answers to question 2A for the open versions of the map skills test

	Inductive	Deductive	Total
Recognizing associations	15	0	8
Classifying; recognizing a single distribution	31	8	20
Other	54	92	72
Total percent	100	100	100
N	26	24	50

The next section of a protocol in the qualitative research shows clearly how a student who does have sufficient geographical procedural knowledge reasons. Marvin answers question 2A (inductive open version):

"I look at all the nuclear reactors, where they are all located. Then you can sort of make a comparison. There's one close to the water, mostly next to a river. Always.

C is the wrong answer, because it's not close to anything. Neither are E and D.... (The interviewer points out that the answer categories are A, B, C and D) A, B, C en D. Then B is the best answer. All you have to do is compare ... most of them are close to water and B is close to the river too".

Two other kinds of answers to question 2A are worth mentioning. In the first place, there are a few students who have difficulty reading a map. An example of this is a student who says for question 2A that location A would be a good place to build a nuclear reactor "Because A is at the sea for cooling-water".

In this case appropriate declarative knowledge was used, but no attention was paid to distances on the map. In the second place there are students who used incorrect arguments for a good choice for locations. An example: "B, because it is flat and easy to build there". This student not only uses incorrect declarative knowledge, but is also unable to find the generalizations on the map and therefore does not have sufficient geographical procedural knowledge.

Using both the research results discussed above and the literature mentioned we can schematically order the reactions of the students to the map skills test we used (see figure 6.8). The ordering runs from 'very desirable' to 'not desirable' student reactions. For geography it is desirable to make combined use of declarative and procedural knowledge (Van Westrhenen 1986).

- | | |
|-----|---|
| I | The student uses both procedural and declarative knowledge |
| a | The student can interpret maps |
| b | The student can analyze maps: |
| | • The student can recognize areal associations (relate) |
| | • The student can recognize spatial distributions or areal differentiations (classify). |
| c | The student can read maps (identify). |
| II | The student only uses declarative knowledge |
| III | The student does not give an answer or just repeats that which is asked or gives an answer which is not relevant to the question. |

Figure 6.8 Categories for student reactions to map skills tests

The research described here aims at finding out how many students score in which category in figure 6.8 and at finding out how students can be stimulated to use not only declarative knowledge but procedural knowledge as well. It is especially the subcategories for procedural knowledge that are of importance to this research. The ability to identify, classify and relate are not subject specific skills. In geography these skills can be aimed at being able to read maps, being able to recognize spatial distributions and areal differentiations and being able to recognize areal associations.

6.7 A few final remarks

The research results described above are only a brief report on a first exploratory investigation. This exploratory investigation is part of the larger research program described in the first five paragraphs of this article.

Combining quantitative research (written tests) with qualitative research (thinking-aloud protocols) seems to be a good method by which to discover those processes that students go through as they analyze issues for which they need to be able to use geographical skills.

The research results described here show that many students have problems with applying procedural knowledge. Most students have little trouble recognizing declarative knowledge. Applying procedural knowledge is quite difficult for most pupils, as has been shown earlier by research programs concerning map analysis skills of pupils in secondary education in the Netherlands (Van der Schee 1985; Van der Schee 1987). It is easier to understand these research results when one realizes that in Dutch geography classes currently almost all of the time is spent on declarative knowledge and that there is hardly any systematic training in procedural knowledge at all. As far as procedural knowledge and map analysis is concerned, Van der Schee concluded in 1988 that there is an unbalanced attention for map skills in the most used Dutch school geography books for the first years of secondary education: "The 280 thematic maps and connected map assignments in these books were classified with help of key concepts which are related to the geographical method. It appeared that a lot of attention was paid to the concept of spatial distribution and less to other geographical key concepts: spatial association, spatial interaction, spatial diffusion and spatial system".

It also appears that many students do not think in spatial or areal patterns but are mainly object-oriented. This conclusion agrees with results obtained in previous research in geography teaching at primary schools. Writing about a Dutch research program, Van Westrhenen (1985, p.50) concluded that geography lessons structured according to geographical disciplinary concepts can help young students to discover the spatial context of their actions and to see more than their own experiences. Dijkink & Elbers (1978, pp.82-88) asked children of different ages to build a city and found a clear relation between the development of geographical concepts and the age of the students: "The level of analysis of younger children is the level of the element..... and some older children are able to think in terms of aggregates".

Qualitative research shows especially clearly that not all students use all of the available information and that map information is certainly not used in the first instance. Schouten-Van Parreren (1989) reports similar findings for her quality research on reading ability and vocabulary acquisition in French (This is part of the larger research project). Just as in our research in geography, the pupils in the French research project have a number of different sources of information. These include previous knowledge, the texts, the pictures and the directions in the assignments. Schouten-Van Parreren has discovered that the weaker students often use just once source of information. Weaker students appear to be unable to use different sources of information systematically to select and integrate informa-

tion: "It is important to help weaker students in using different sources of information systematically. If no help is given, the differences between students will grow, even when pictures are offered together with texts, what is often done to help the weaker students!" (Schouten-Van Parreren 1989, p.35). A similar pattern can be recognized in the thinking-aloud protocols for the qualitative research for geography.

It is not easy to design good geographical test questions. In addition to geographical complexity we have to deal with the wording of the question, the cartographic presentation and the nature of the phenomena displayed on the map. "What is the effect on the map skills of students of the scale of the map and using fictitious or real maps? Is there a relation between map skills and the amount of relevant and irrelevant information on the map?" (Van der Schee 1985, p.425).

Consulting seventeen experts on geographical test question design has provided useful suggestions for improving the map skills test. Using this information together with the results of the initial explorative research described here we will need to refine and expand the map skills test. Further explorative research with an improved research instrument is needed in order to be able to make reliable statements about the skills students use in dealing with different aspects of geographical procedural knowledge.

References

- BEAUJEU-GARNIER, J. (1976), *Methods and perspectives in geography*. London: Longman.
- BOEKAERTS, M. (1983), Enkele kanttekeningen bij het interactionistische motivatiemodel. In: Th. Bergen & E. Roede - *Motivatie gemeten?* Harlingen: SVO.
- BOEKAERTS, M. (1987), *Psychologie van de leerling en het leerproces*. Nijmegen: Dekker & Van de Vegt.
- BOONMAN, J.H. & W.A.M. KOK (1982), *Het verwerven van kennis*. Losbladig Onderwijskundig Lexicon, PO 4240, pp. 1-18.
- BREUKER, J.A. & J.J. ELSHOUT & M.W. VAN SOMEREN & B.J. WIELINGA (1986), Hardop-denken en protokolanalyse. *Tijdschrift voor Onderwijsresearch* pp. 241-254.
- COMMISSIE HERZIENING EINDTERMEN (1990), *Advies kerndoelen voor de basisvorming in basisonderwijs en voortgezet onderwijs*. Den Haag: DOP.
- COOK, Th.D. & D.T. CAMPBELL (1979), *Quasi-experimentation; design and analysis issues for field settings*. Chicago: Rand McNally.
- DIJKINK, G.J. & E. ELBERS (1978), *Over de geografische representatie bij kinderen*. Amsterdam: Vrije Universiteit.
- HOEKVELD, G.A. (1989), Over de methodiek van de regionale geografie. In: G.A. Hoekveld & G. Schoenmaker & J. van Westrhenen, eds., *Wijkende grenzen*. Amsterdam: KNAG/Vrije Universiteit, pp. 19-44. (Netherlands Geographical Studies 103)
- HOEKVELD-MEIJER, G. & G. SCHUTTE (1988), *Aardrijkskunde gebiedsonderwijs; tekst en uitleg bij het lezen, denken en leren over gebieden en verschijnselen in gebieden*. Amsterdam: KNAG/Vrije Universiteit. (Netherlands Geographical Studies 82)
- KLERK, L.F.W. de (1983), *Onderwijspsychologie*. Deventer: Van Loghum Slaterus.
- KNOERS, A.M.P. (1973), *Leren en ontwikkeling, leerpsychologie ten behoeve van het onderwijs*. Assen: Van Gorcum.
- PARREREN, C.F. VAN (1974), Het functioneren van leerresultaten. In: C.F. van Parreren & J. Peeck. *Informatie over leren en onderwijzen*. Groningen: Wolters-Noordhoff. pp. 114-131.

- POSNER, G.J. & K.A. STRIKE (1976), A categorization scheme for principles of sequencing content. *Review of Educational Research*, 46, pp. 665-690.
- SCHEE, J.A. VAN DER (1985), Kaarten geven te denken. In: *Geografisch Tijdschrift XIX* pp. 418-426.
- SCHEE, J.A. VAN DER (1987), Kijk op kaarten. Amsterdam: Vrije Universiteit (Netherlands Geographical Studies 36).
- SCHEE, J.A. VAN DER (1988), Map and map skills. In: R. Gerber & J. Lidstone, eds., *Skills in Geographical Education, Symposium Proceedings*. Brisbane: IGU, 2, pp. 685-694.
- SCHOUTEN- VAN PARREREN, C. (1989), Een handelingspsychologische analyse van verschillen tussen leerlingen bij het leren van vreemde talen. In: *Handelingen*, 3, pp. 18-36.
- SPRUIT, P. T. (1974), De geografische beschouwingswijze. In: *Geografie en basisonderwijs*. Hoevelaken: LPC, pp. 7-33.
- WESTRHENEN, J. VAN (1976), Aardrijkskundige onderwijsdoelen. Meppel: Krips.
- WESTRHENEN, J. VAN (1985), De bijdrage van vakken aan de leergebieden 'mens en maatschappij' en 'natuur'. In: G.M.A. Verschuren, ed., *Vakkengesplitst onderwijs - Vakkengericht onderwijs*. Amsterdam: KNAG, pp. 31-59.
- WESTRHENEN, J. VAN (1986), Basisvorming in het voortgezet onderwijs als kader voor de aardrijkskunde. In: *Geografisch Tijdschrift XX*, no. 4, pp. 293-304.
- WESTRHENEN, J. VAN (1989a), Regionale kennisoverdracht in nieuw perspectief. In: G.A. Hoekveld, G. Schoenmaker & J. van Westrhenen, eds., *Wijkende grenzen*. Amsterdam: KNAG/Vrije Universiteit, pp. 45-70 (Netherlands Geographical Studies 103).
- WESTRHENEN, J. VAN (1989b), Het ontwerpen van aardrijkskundige leerstof en didactische differentiatie. In: J.A. van der Schee & H. van Helden, eds., *Educatief Perspectief*. Amsterdam: Bureau Lerarenopleiding VU, pp. 57-82.

7 HOW DO STUDENTS IN HIGHER EDUCATION PERCEIVE MAPS?

Rod Gerber

7.1 The problem

The design of maps has fascinated people ever since these graphic forms of communication have been drawn. The problems of converting geographic information from a three-dimensional form to a two-dimensional form have been a major pre-occupation for cartographers for hundreds of years.

Cartographic theorists such as Raisz (1938), Robinson (1960), Muehrcke (1978) and Dent (1988) have all emphasized the challenges of map projections. Psychologists such as Piaget & Inhelder (1956), Paivio (1971), Haber & Hershenson (1974) and Hochberg (1962) have drawn attention to the perceptual challenges that confront people as they perceive spatial arrangements and relationships. In the 1980's, cartographic researchers including Bertin (1981, 1983), Bos (1984) and Gerber (1980; 1987; 1988) have refocused the map design challenge to highlight the map user in the cartographic communication process. They argue that, based on the perception and understanding of graphic symbols by these users, map designers develop useful and effective cartographic representations.

One critical challenge in the design of user-driven cartographics is the development of a definite visual hierarchy on each map. As Bertin (1981) established, the variables that underpin such a hierarchy consist of color, shape, size and orientation. This challenge lay at the foundation of the study reported in this paper. Specifically, this study investigated the perception and understanding of visual hierarchies on maps based on color in point, linear and areal symbols and on map type (the names on maps).

7.2 Design of the study

Investigations such as the one reported in this article often experience difficulties because of the inability to control the range of variables that impinge on the study. It is for this reason that a block design was selected and the participants in the study were generally homogeneous.

The blocks in the design are summarized in table 7.1 below. Basically, they consist of three sets of four maps that have been designed around (a) point symbols, (b) linear symbols, and (c) areal symbols.

On each map, fifty-five names of places were included. Each set of four maps was designed on the following basis:

- (a) a visual hierarchy based on color, plus a visual hierarchy based on map type;
- (b) a visual hierarchy based on color; no visual hierarchy based on map type;
- (c) no visual hierarchy based on color, but a visual hierarchy based on map type;
- (d) no visual hierarchies based on either color or map type.

Table 7.1 Block design used in visual hierarchies study

Map number	Colour hierarchy	Type hierarchy
1 (Point symbols)	Yes	Yes
2 (Point symbols)	Yes	No
3 (Point symbols)	No	No
4 (Point symbols)	No	Yes
5 (Linear symbols)	Yes	Yes
6 (Linear symbols)	Yes	No
7 (Linear symbols)	No	No
8 (Linear symbols)	No	Yes
9 (Areal symbols)	Yes	Yes
10 (Areal symbols)	Yes	No
11 (Areal symbols)	No	No
12 (Areal symbols)	No	Yes

Such a design enabled the investigator to work with a relatively homogeneous group and to control the use of color and type. However, it was not possible to control a series of other user-based variables including home environment, previous experience using maps, drawing ability, spatial ability and age that the author had found previously (Gerber 1980) were significant influences on users' competence in map language.

This design dictated the use of a strongly positivistic approach to the investigation - an approach characterized by the use of standard quantitative research methods that were supported by a range of open-ended responses.

7.3 Instruments used in the study

As mentioned in table 7.1, the instruments used in the study consisted of twelve original thematic maps designed by a cartographer, using a tetrax computer, under instructions from the author. These maps (see figures 7.1-7.12) were designed based on the following principles:

- (a) each map would occupy the same area drawn at the same scale with a similar orientation;
- (b) the color hierarchy used on the maps is based on the principle of the darker the color and the larger the symbol, the more important/larger/significant the feature represented;

- (c) the type hierarchy used on the maps consisted of the larger, thicker and bolder the type, the more important the place name.

The resulting hierarchy of *point* symbols consisted of:

- (a) a dark red, 4 mm square for settlements over one million people;
- (b) an orange, circular shape, 4 mm in diameter, for settlements containing 500 000-1 000 000 people;
- (c) a black, stippled square, 2 mm wide, for settlements between 100 000 and 499 999 people; and
- (d) fine, red, stippled circle, 2 mm in diameter for settlements containing less than 100 000 people.

The hierarchy of *linear* symbols consisted of:

- (a) red-black lines, 4 mm wide, for international highways;
- (b) red lines, 2 mm wide, for national roads;
- (c) black stippled lines, 1 mm wide, for state roads; and
- (d) fine, red dotted lines for local roads.

The hierarchy of *areal* symbols consisted of:

- (a) black shading for more than 2 000 people per square kilometer;
- (b) orange shading for 1 000-1 999 people per square kilometer;
- (c) gold shading for 100-999 people per square kilometer; and
- (d) pale yellow shading for less than 100 people per square kilometer.

The hierarchy of *map type* used consisted of:

- (a) 10 point Univers Medium Bold Caps Condensed type;
- (b) 10 point Univers Medium Bold type;
- (c) 8 point Univers Medium Bold type; and
- (d) 6 point Univers Medium Bold type.

Six point Univers medium bold type was used where there was no type hierarchy.

These instruments formed the basis for twenty tasks that were designed to ascertain the participants' perception and understanding of the presence of visual hierarchies on the maps. For each set of four maps, the following sequence of tasks was developed:

- (a) three questions on understanding the geographical information on each of the maps;
- (b) one question about the reason for identifying different levels of data on each map;
- (c) one question to identify the symbols that attracted attention on each map;
- (d) one question to identify on which of the four maps in the set the symbols were clearest;
- (e) one question to establish the reason for choosing the map in (d) above.

For all of the maps, the participants were asked to think of the combination of point, linear, areal symbols and map type, in order to choose:

- (a) the clearest representations of point, linear and areal symbols; and
- (b) to indicate reasons for these choices.

The data on these maps and the twenty tasks were evaluated in a pilot study with ten cartography students and minor adjustments were made before actual study was implemented.

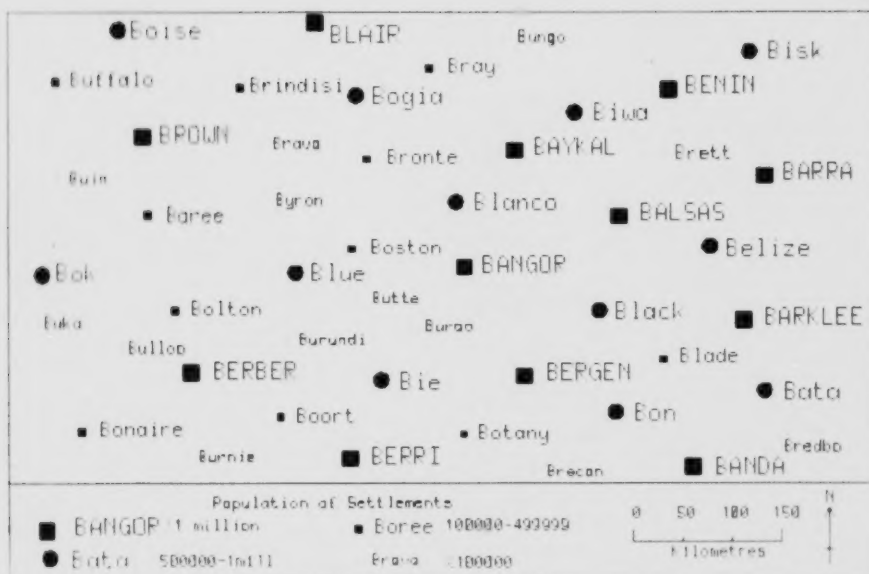


Figure 7.1 Map 1 - Settlements in Bentinck

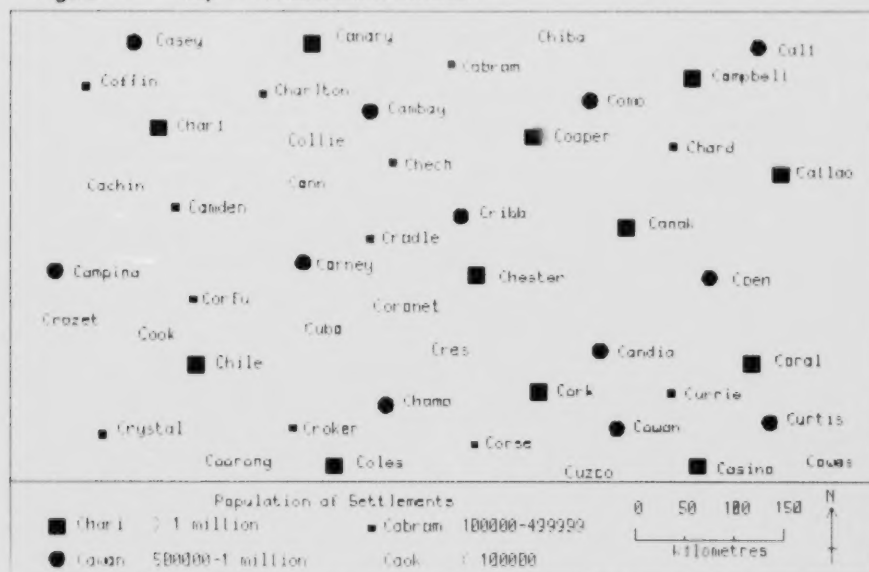


Figure 7.2 Map 2 - Settlements in Comorin

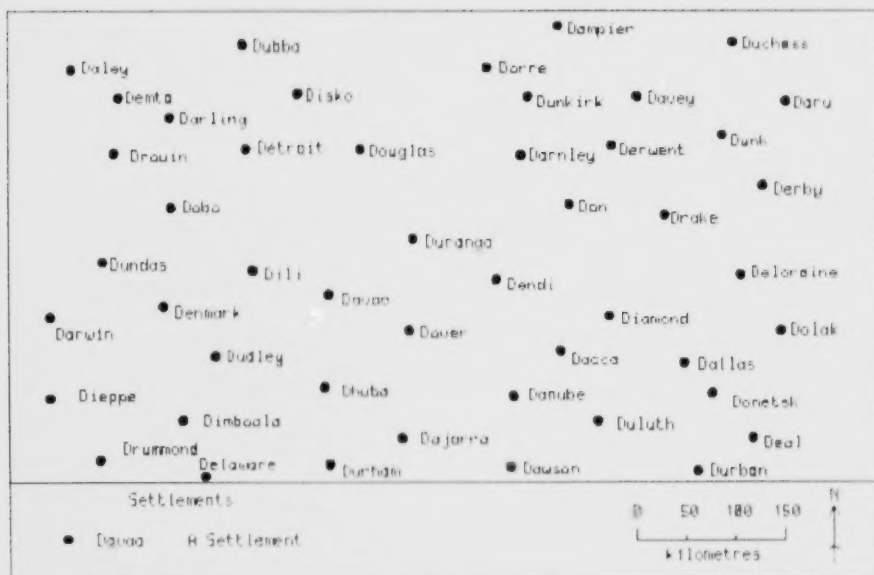


Figure 7.3 Map 3 - Settlements in Dargaville



Figure 7.4 Map 4 - Settlements in Eaglehawk

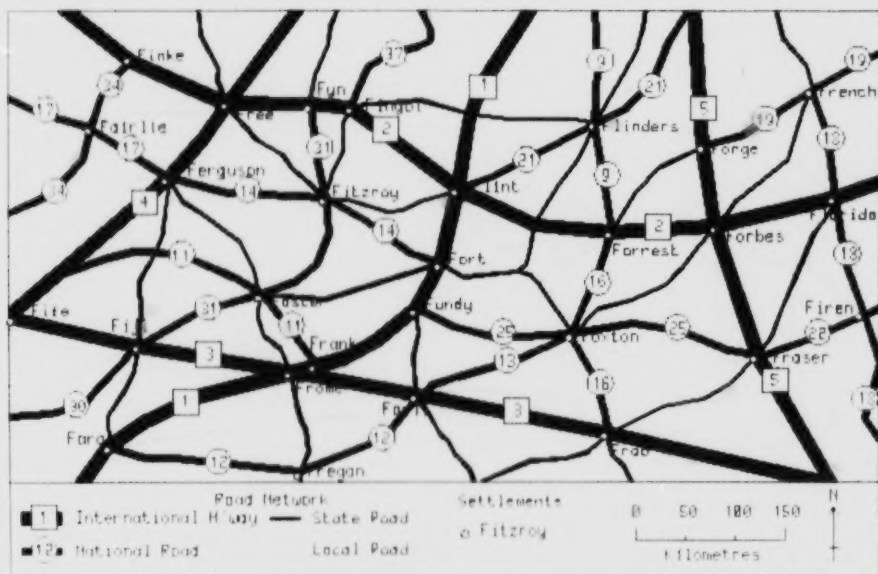


Figure 7.5 Map 5 - Road network in Ferrara

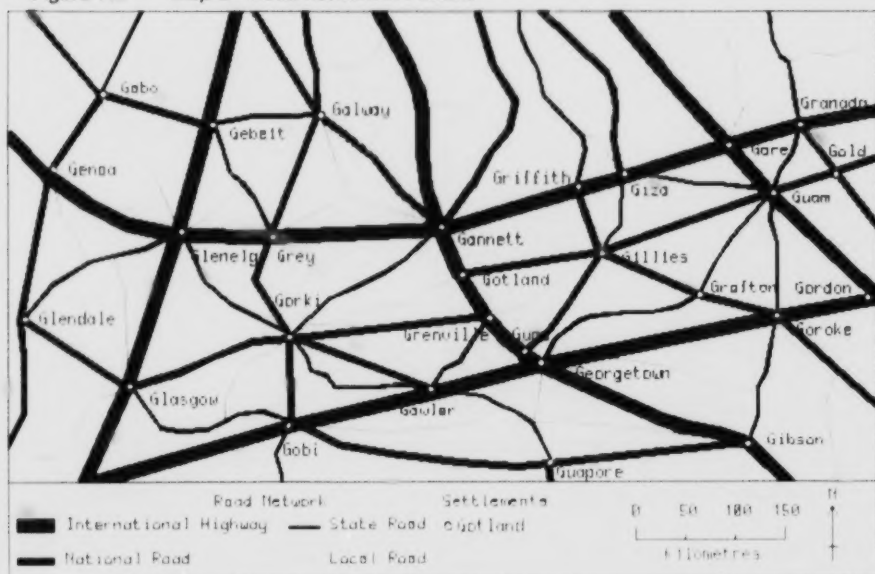


Figure 7.6 Map 6 - Road network in Good Hope

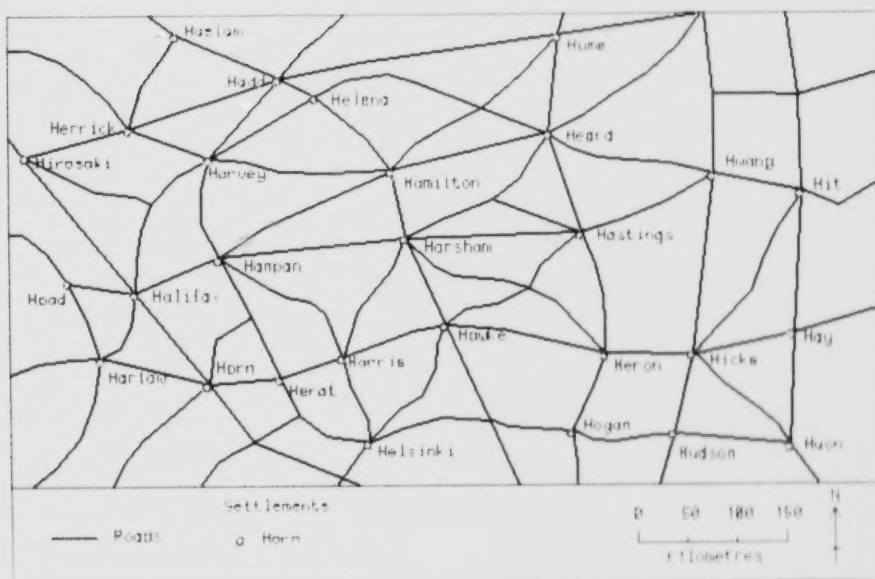


Figure 7.7 Map 7 - Road network in Heathecote

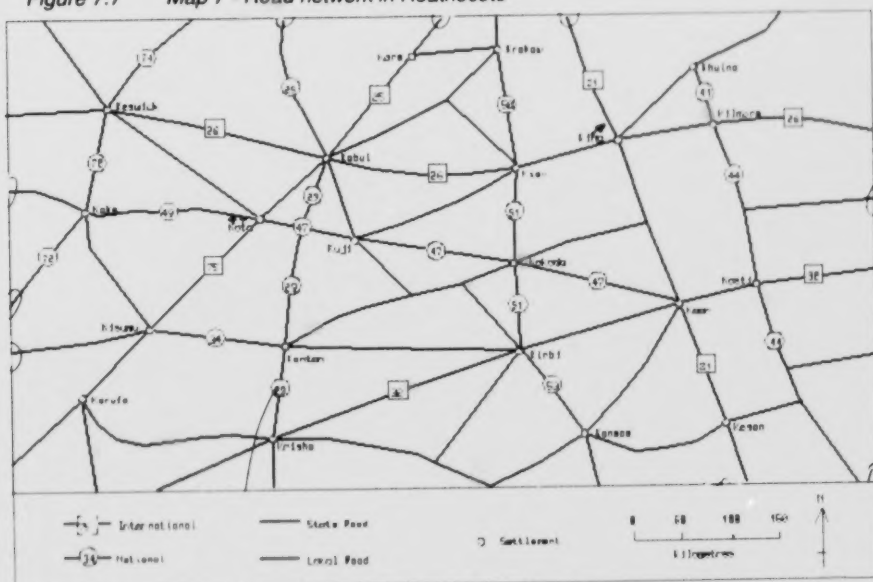


Figure 7.8 **Map 8 - Road network in Kandrian**

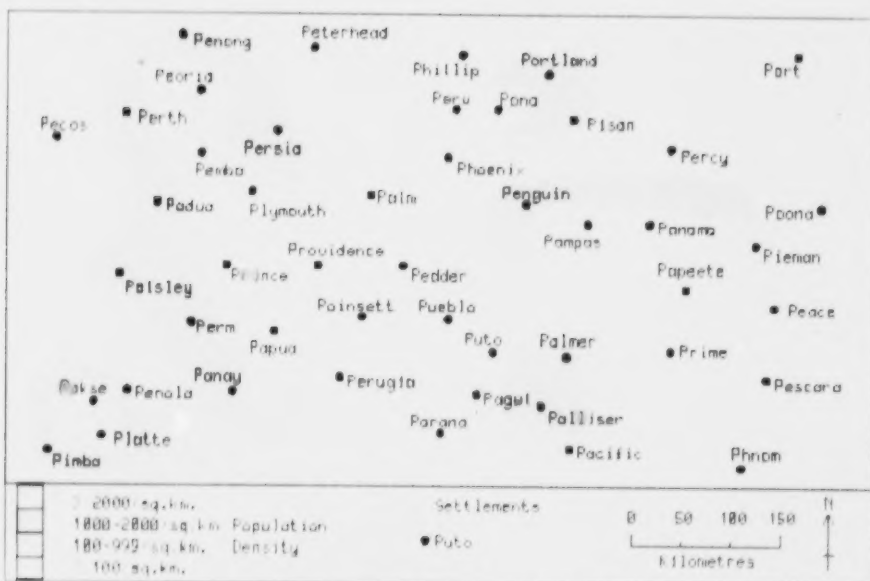


Figure 7.11 Map 11 - Population density in Parinas

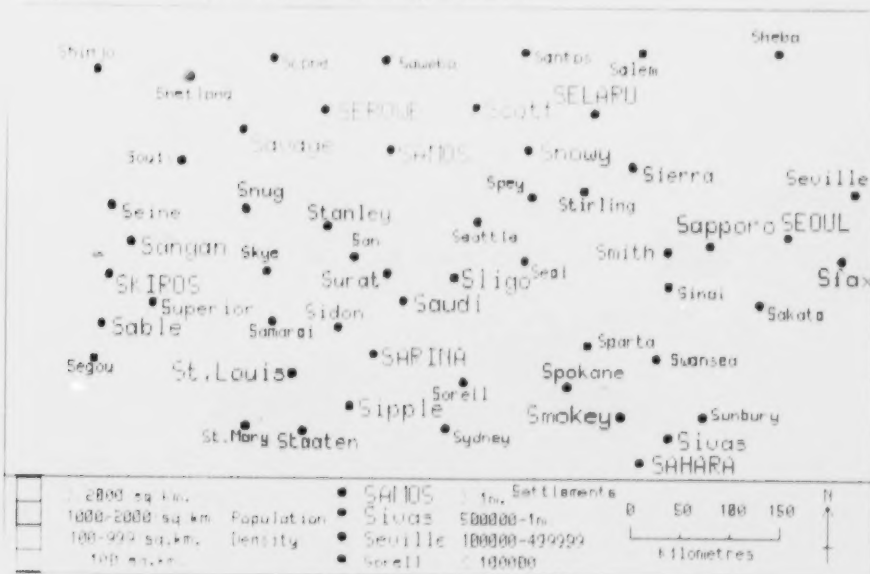


Figure 7.12 Map 12 - Population density in Sherpa

7.4 The participants in the study

Eighty-one students who were majoring in geography and/or cartography in two Australian universities were voluntary participants in this investigation. All of these students had studied some cartography or geography in an university. By choosing such a group of participants, it was expected that some control could be placed over the uncontrolled variables in the study. For example, it was assumed that such a group would possess similar levels of intellect, interests in geographical information and skills in map-reading. The extent to which these assumptions are true was not tested during the study.

7.5 Analyses of the data

The data which were obtained by having the eighty-one students complete the range of tasks mentioned above with the twelve maps, was analyzed in the following ways:

- (a) the calculation of simple frequency counts for the students' preferences of selected symbols on particular thematic maps. These frequencies are expressed as percentages;
- (b) the calculation of the frequency of competence (i.e. 100 per cent accuracy) in the students' understanding of the use of visual hierarchies in the variables of color and type in thematic maps;
- (c) the use of simple frequency counts to quantify the reasons that students had for telling the differences between the symbols used in these visual hierarchies;
- (d) the calculation of Spearman Correlation Co-efficients to investigate the existence of any significant relationships between the students' preferences and their understandings of the visual hierarchies in the three sets of maps.

The calculations of students' preferences, their understandings and the Spearman Correlation Co-efficients were completed using programs from the SAS package. The frequency counts of their reasons were completed manually.

7.6 Point symbols

The students' preferences and understanding of the visual hierarchies in the four maps associated with point symbols and type are summarized in table 7.2 below. They indicate that 93.7 per cent of the participants preferred map 1 which contained visual hierarchies in both color and type. The extent of the students' understanding of the geographical information based on the extent to which visual hierarchies were used, varied considerably across the four maps. Best understanding was achieved in map 1 (96.2 per cent of the students demonstrated complete accuracy in their responses) in which both visual hierarchies were present and in map 2 (91.2 per cent of the students demonstrated complete accuracy in their responses) where there was a visual hierarchy in only color. Least understanding was displayed with the maps that contained no visual hierarchy in the color of the symbols.

Table 7.2 Students' preferences and understanding of visual hierarchies on thematic maps using point symbols (per cent)

Preference of map		Total accuracy in understanding	
Map 1	93.7	Map 1	96.2
		Map 2	91.2
		Map 3	83.7
		Map 4	61.2

It was interesting to note that the understanding was at its lowest on the map that contained no visual hierarchy based on color, but did contain a type hierarchy. The students' reasons for identifying the different sizes of the settlements on the four maps are summarized in table 7.3 below.

Table 7.3 Frequencies of students' reasons for identifying different sizes of settlements on maps 1-4

	Map 1	Map 2	Map 3	Map 4
Different sizes of symbols	43	30	0	0
Different colours of symbols	42	41	0	0
Size of the type	10	4	3	67
Shapes of the symbols	12	24	0	3
The legend	8	21	0	3
Don't know	0	0	63	0

Their reasons indicate that the majority of them are well aware of the visual hierarchies that have been used on the four maps that are based on point symbols. Where two hierarchies are evident, the students showed a definite preference for the hierarchy built around color. Where the hierarchy was built around either color or type, the students had little difficulty appreciating this aspect of design. Similarly, where there was no intended visual hierarchy, the students also appreciated that fact.

Table 7.4 Frequencies of students' selection of the most visible symbols on maps 1-4

	Map 1	Map 2	Map 3	Map 4
Red squares and orange circles	70	76	0	0
Red circles	0	0	41	21
Large map type	6	0	0	38
Cannot tell	0	0	36	9

When the students were asked to highlight the symbols on each map that attracted their attention, they invariably focussed on the symbols or type that were intended to be the most visible in the hierarchy (see table 7.4). For maps 1 and 2, these were the red squares or orange circles and for map 4, it was the size of the lettering supported by the red circles. For map 3, where there was no intended hierarchy, the students either couldn't decide or reverted to the red circles.

The final analysis concerning this set of maps concerned the calculation of Spearman's Correlation Co-efficient to determine if there were any significant relationships between the preferred map based on point symbols and the students' understanding of the visual hierarchies in the geographical information. The results of this calculation are summarized in table 7.5. They indicate that the only significant relationship was a moderate positive correlation between the preferred map and the understanding shown for the information on map 4 which contained only a hierarchy of map type. The precise explanation for this correlation is not clear. It may mean that those students who prefer visual hierarchies based on map type have a better understanding of this hierarchy than if the hierarchy is based on color.

Table 7.5 Spearman's correlation co-efficients between students' perceptions and understandings of visual hierarchies on maps 1-4

	Spearman correlation co-efficient	Probability
Map 1	0.010	0.928
Map 2	0.015	0.888
Map 3	0.022	0.842
Map 4	0.299	0.007*

* Significant at 0.05 level

7.7 Linear symbols

The students' preferences and understanding of the visual hierarchies in the four maps associated with linear symbols and type are summarized in table 7.6 below. They indicate that 83.7 per cent of the participants preferred map 5 which contained visual hierarchies in both color and type. The extent of the students' understanding of the geographical information, based on the extent to which visual hierarchies were used varied considerably across the four maps. Best understanding was achieved on map 7 where there were no visual hierarchies used (83.7 per cent achieve total accuracy) and on map 6 where there was a hierarchy of linear symbols only (82.5 per cent achieved total accuracy). Least understanding was achieved on map 8 (51.2 per cent achieved total accuracy) which contained only a visual hierarchy of map type. The relatively low performance on map 5 which contained visual hierarchies in both color and map type is interesting.

Table 7.6 *Students' preferences and understanding of visual hierarchies on thematic maps using linear symbols (per cent)*

Preference of map		Total accuracy in understanding	
Map 5	83.7	Map 5	62.5
		Map 6	82.5
		Map 7	83.7
		Map 8	51.2

The students' reasons for identifying the different levels of importance for the roads on the four maps are summarized in table 7.7 below.

Table 7.7 *Frequencies of students' reasons for identifying the level of importance on roads on maps 5-8*

	Map 5	Map 6	Map 7	Map 8
Varying widths of road symbols	44	69	0	0
Different colours of road symbols	39	65	0	0
Road identifiers (numbers)	8	0	0	58
Shape of the symbols	1	7	4	14
The legend	12	6	0	3
Don't know	0	0	63	5

Their reasons indicate that the majority of them are well aware of the visual hierarchies that have been used on the four maps that are based on linear symbols. Where the two hierarchies are evident, the students showed a definite preference for the hierarchy built around color. This has probably occurred because the labels used for the road symbols consisted of highway identifiers which many students assumed were a part of the linear symbols. A stronger set of reasons were offered on map 6 where the only hierarchy was based on the color of the road symbols. Students had great difficulty offering reasons on map 7 which contained no intended visual hierarchy. In map 8, the visual hierarchy built around the highway markers was definitely understood.

When the students were asked to highlight the symbols on each map that attracted their attention, they invariably focussed on the symbols or type that were intended to be most visible in the respective hierarchies (see table 7.8). For map 5, these consisted of the thick red lines and the highway markers. The thick red lines on map 6 were most noticeable. On map 7 which contained no hierarchy, the thin red lines were the only noticeable symbol. A considerable number of students could not find any attractive symbols. On map 8, the highway markers (i.e. the type hierarchy) were the most obvious symbols to the students.

Table 7.8 *Frequencies of students' selection of the most visible symbols on maps 5-8*

	Map 5	Map 6	Map 7	Map 8
Red lines	56	57	29	6
Highway markers	18	13	0	42
Numbers on highways	11	0	0	16
Road network	2	0	9	2
None	0	0	22	10

The final analysis concerning this set of maps concerned the calculation of Spearman's Correlation Co-efficient to determine if there were any significant relationships between the preferred map based on linear symbols and the students' understanding of the visual hierarchies in the geographical information. The results of this calculation are summarized in table 7.9. They indicate that there are no significant relationships between map preference and understanding of visual hierarchies associated with color in linear symbols and the use of type or numbers on thematic maps.

Table 7.9 *Spearman's correlation co-efficients between students' perceptions and understandings of visual hierarchies on maps 5-8*

	Spearman correlation co-efficient	Probability
Map 5	0.074	0.512
Map 6	0.045	0.685
Map 7	-0.034	0.760
Map 8	0.127	0.260

7.8 Areal symbols

The students' preferences and understanding of the visual hierarchies in the four maps associated with areal symbols and type are summarized in table 7.10 below. They indicate that 53.7 per cent of the participants preferred map 12 which contained a visual hierarchy built around the principle of "the higher the amount, the lighter the color of the areal symbol" and which contained a hierarchy of map type, while 37.5 per cent of the participants preferred map 9 which contained a hierarchy of color based on the principle, "the higher the amount, the darker the color of the areal symbol" and which contained a hierarchy of map type. The extent of the students' understanding of the geographical information did not vary greatly across the four maps. In fact, the range was only from 72.5 per cent total accuracy on map 11 to 81.2 per cent total accuracy on map 9.

Table 7.10 *Students' preferences and understanding of visual hierarchies on thematic maps using areal symbols (per cent)*

Preference of map		Total accuracy in understanding	
Map 9	37,5	Map 9	81.2
		Map 10	78.7
		Map 11	72.5
		Map 12	75.0
Map 12	53.7		

The students' reasons for identifying the different populations of the settlements and the different population densities of the areas on the four maps are summarized in table 7.11 below.

Table 7.11 *Frequencies of students' reasons for identifying the populations of settlements and the different population densities of areas on maps 9-12*

	Map 9	Map 10	Map 11	Map 12
Different colours	66	50	0	61
Size of the lettering	45	0	0	55
Legend	6	3	0	9
Cannot tell	0	21	81	0

Their reasons indicate that the majority of them are well aware of the visual hierarchies that have been used on the four maps that are based on areal symbols. For the hierarchies of areal symbols based on either principle, "the higher the lighter" (map 12) or "the higher the darker" (map 9), the students were very aware of both principles. They did not favor one principle over the other. Where a type hierarchy was included (maps 9 and 12) the students were well aware of its presence on the maps. However, not as many students emphasized the type hierarchy over the color hierarchy. It is noticeable that for the thematic map which contained neither hierarchy, the students were unanimous in being unable to make definite decisions on the selection of geographical information symbols.

When the students were asked to highlight the symbols on each map that attracted their attention, they invariably focussed on the symbols or type that were intended to be the most visible in the respective hierarchies (see table 7.12). For maps 9 and 10, these consisted of the dark shading. In map 9, the large-sized map type was noted. Even in map 11 where there was no intended color hierarchy, the students identified the different colors for the areal symbols. In map 12, both hierarchies were identified in the yellow shading and the large map type.

Table 7.12 *Frequencies of students' selection of the most visible symbols on maps 9-12*

	Map 9	Map 10	Map 11	Map 12
Dark colours	56	57	0	0
Large map type	21	0	0	33
Pastel colours	0	0	48	58
None	3	3	7	0

The final analysis concerning this set of maps concerned the calculation of Spearman's Correlation Co-efficient to determine if there were any significant relationships between the preferred map based on areal symbols and the students' understanding of the visual hierarchies in the geographical information. The results of this calculation are summarized in table 7.13. They indicate that the only significant relationship is a moderate positive correlation between preference and understanding in map 12 which contained a hierarchy of map type and a hierarchy of color built on the principle of "the higher the lighter".

Table 7.13 *Spearman's correlation co-efficients between students' perceptions and understandings of visual hierarchies on maps 9-12*

	Spearman correlation co-efficient	Probability
Map 9	0.158	0.159
Map 10	-0.107	0.343
Map 11	-0.072	0.523
Map 12	0.296	0.007*

* Significant at the 0.05 level

7.9 Combination of symbols

After the above tasks were completed, the students were asked to consider a combination of the types of symbols from all of the maps, to arrive at a preferred range of symbols. The summary of their preferences is presented in table 7.14. It reveals that the students generally supported their above decisions with point symbols being derived from map 1, linear symbols from map 5, areal symbols from either map 12 or map 9, and type from map 1 or map 12. Most consensus was established over point and linear symbols and least consensus over map type.

Table 7.14 Selected combination of map symbols from maps 1-12

	Map	Percentage support
Point symbols	1	88.7
Linear symbols	5	65.0
Areal symbols	12	43.7
	9	40.0
Map type	1	30.0
	12	27.5

A range of reasons was suggested by the students for this selection. These reasons are summarized in table 7.15 below. These reasons include: colors used, the size of the symbols, the variety of the type used and the combination of symbols and type. These reasons are closely allied to the two types of hierarchy that were used on the twelve maps.

Table 7.15 Reasons for the selection of the combination of map symbols from maps 1-12

	Number
Colours of the symbols	20
Clarity of the symbols	19
Combination of colour and type	17
Size of symbols	13
Type variation	10
Clear legend	6

7.10 Discussion

The results of this study revealed interesting variations in the students' perception and understanding across the different types of map symbols and map type. The variables of color and map type do provide a very sound basis for understanding how people appreciate visual hierarchies. However, this appreciation does not occur equally across the main types of map symbols.

For point symbols, this study indicates that higher education geography/cartography students have a reasonable linkage between their perception and understanding of thematic maps based on point symbols that are constructed on a hierarchy of color whether or not there is a hierarchy of map type. Occasional variations in the shape of these symbols appear to enhance the students' differentiation of the hierarchy of point symbols. Despite this linkage, there was only a significant relationship between perception and understanding of information on the map that contained a type hierarchy and no hierarchy of point symbols. The expressions of reasons for the perceived hierarchies indicate that most of the students do, in fact, understand the bases for the hierarchies even if they do not apply them in their map reading tasks. It is reasonable to conclude that students in the study did grasp the concept of a visual hierarchy of point symbols based on color and supported by shape variations.

The students' perception and understanding of linear symbols that are organized on a visual hierarchy based on color are less clear than for point symbols. Firstly, it is noted that the differences between the students' perceptions and their understanding of information based on linear symbols were considerable. While there was a definite preference for a particular map containing the intended visual hierarchy, this was not matched by the students' understanding of the information. Despite these performances, the students were still able to nominate the bases for the hierarchies of linear symbols and type. There were no significant relationships between the students' perception and understanding of the visual hierarchies. These findings suggest that the students were able to understand the use of color in the development of a visual hierarchy for linear symbols. However, such an hierarchy was not seen to be important in the use of these symbols on thematic maps.

If linear symbols provided a mixed response for the students in the study, the perception and understanding of visual hierarchies based on color in areal symbols is still less clear. An obvious explanation for this conclusion is that areal symbols can be constructed using the principles of "the higher the lighter" or "the higher the darker". This study indicated that either principle was acceptable for a perceptual viewpoint as long as there was a type hierarchy included. However, there was no evidence of better understanding using areal symbols whether there was a visual hierarchy present or not. It is worth noting that there was a significant relationship between the students' perception and understanding of areal symbols based on the principle of "the higher the lighter". Again, students were aware of the main attributes of the visual hierarchy based on color for areal symbols.

Students' perception of a type hierarchy seems to be clearest on maps containing point symbols based on a strong color-based visual hierarchy or maps containing areal symbols based on the principle of "the higher the lighter". This second alternative conforms to the principle adopted by many cartographers in their design of layer-tinted maps.

These conclusions lead to a final generalization that it is very probable that visual hierarchies in cartographic symbols are valuable in both the perception and understanding of the geographical information on maps. Cartographers, however, should

be aware that not all visual hierarchies based on color or type will succeed as well as the designer intended. However, the resulting perceptual qualities associated with these hierarchies do often assist users in deriving meaning from thematic maps. They should be used more systematically on all types of thematic maps.

References

- BERTIN, J. (1981), *Graphics and graphic information processing*. Berlin: De Gruyter.
- BERTIN, J. (1983), *Semiology of graphics*. Wisconsin, U.S.A.: University of Wisconsin.
- BOS, E. (1974), Systematic symbol design in cartographic education. *ITC Journal* 1, pp. 20-28.
- DENT, B. (1988), *Principles of thematic map design*. U.S.A.: Addison-Wesley.
- GERBER, R. (1980), Development of competence and performance in cartographic language by children at the concrete level of map-reasoning. Brisbane: University of Queensland.
- GERBER, R. (1987), Educating cartographers on the efficient representation of geographical information. Paper presented to the Carto Ed Conference. Melbourne: RMIT.
- GERBER, R. (1988), Seeing things clearly on maps. Paper presented to the National Conference of the Australian Institute of Cartographers. Sydney.
- HABER, R. & M. HERSHENSON (1974), *The psychology of visual perception*. London: Holt, Rinehart and Winston.
- HOCHBERG, J. (1962), The psychophysics of pictorial perception. *AV Communication Review*, 10, Supp. no. 5, pp. 22-54.
- MUEHRCKE, P. (1978), *Map use: reading, analysis, and interpretation*.
- PAIVIO, A. (1971), *Imagery and verbal processes*. New York: Holt, Rinehart and Winston.
- PIAGET, J. & B. INHELDER (1956), *The child's conception of space*. London: Routledge and Kegan Paul Ltd.
- RAISZ, E. (1938), *General cartography*. New York: McGraw-Hill.
- ROBINSON, A. (1960), *Elements of cartography*, second edition. New York: Wiley.

8 GROUP LEARNING IN MAP WORK

Betty Lai-ling Yau, Kam-fai Wong and Lai-sim Ma

8.1 Introduction

In Hong Kong, with the introduction of compulsory education up to age 15 or form 3 (grade 9) in 1978 but with the maintenance of an academically oriented curriculum, many students experience learning difficulties in school. The large classes make it difficult for teachers to give individual students the attention and help they need. Because education is by far the most important means of upward social mobility in Hong Kong, keen competition makes students' plight even worse. The research described in this paper is an attempt to apply the group learning strategy to help students to master map work knowledge and skills in junior classes in a secondary school in Hong Kong.

Map work is an important element in geography. It requires abstract thinking and mathematical calculation skills which are sequentially built up. In Hong Kong, students learn "scale" and "direction" in form 1, "contours" and "contour features" in form 2 and so on, culminating in "map interpretation" in form 5. However, students find it difficult to master the subject, and they repeat the same mistakes over and over in spite of revision. Reports from the last 10 years' Certificate of Secondary Education Examination (grade 11) in geography indicate that map work results were poor in 7 years out of 10, fairly good in only 2 years and quite good in one year. Comments like "students made mistakes in calculation" (vertical exaggeration, gradient and area) appeared in 8 years out of 10, "poor correlation of photographs with maps" appeared in 4 years out of 10, "poor analytical thinking", "poor identification of contour features from topographical maps" and "poor knowledge of conventional signs", each appeared in 2 out of 10 years and "failure to orientate map" in one year (Hong Kong Examinations Authority 1980-89).

8.2 The mastery learning strategy

One of the factors which influenced the development of mastery learning was a 1963 article by Carroll entitled "A model for school learning". Carroll (1963) proposed that if each student was allowed the time needed to learn a subject to some criterion level, and if the student spent that time appropriately, then the student would probably attain the specific level of achievement. Thus the degree of learning attained by the student can be expressed by the following equation:

$$\text{Degree of learning} = f(\text{time spent} / \text{time needed})$$

Carroll further identified the factors that he believed influenced the time spent and

the time needed in learning. Time spent was determined by a learner's (1) perseverance and (2) opportunity to learn. Perseverance is simply the amount of time a student is willing to spend actively engaged in learning. Opportunity to learn is the classroom time allotted to the learning. The time needed was determined by (3) students' aptitude for the subject, (4) quality of instruction and (5) student's ability to understand the instruction. Thus, the equation becomes:

$$\text{Degree of learning} = f((p \times o) / (s \times q \times a))$$

where
 p = perseverance
 o = opportunity to learn
 s = students' aptitude
 q = quality of instruction
 a = students' ability to understand

Bloom (1968) was impressed by Carroll's idea which not only identified the factors which influenced learning in the school setting, it also offered a rather optimistic view that students differed in terms of the time (rather than aptitude) required for learning to reach a mastery level (80-90%) and given that amount of time needed, ideally close to 80% students would attain mastery level.

The preparation processes of the mastery learning strategy are: (1) specify learning objectives and mastery level, (2) decide content, (3) divide and sequence materials into small learning sub-units, (4) devise appropriate teaching strategies (5) set diagnostic test A, (6) prepare corrective activities, (7) prepare enrichment activities, (8) set diagnostic test B and (9) prepare final test (Guskey, 1985).

The teaching processes for the mastery learning strategy are (figure 8.1): (1) specify learning objectives to students, (2) teach unit 1, (3) apply diagnostic test A of unit 1, (4) do corrective/enrichment activities, (5) apply diagnostic test B, (6) apply final test, (7) teach unit 2 and (8) repeat processes (3) to (6) for unit 2.

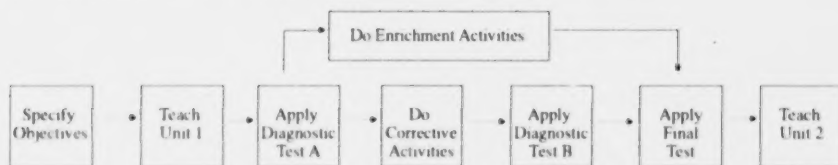


Figure 8.1 Teaching processes of the mastery learning strategy

Two special features of the mastery learning strategy are diagnostic tests and corrective activities. After teaching each small unit, diagnostic tests are used to check learning progress. Diagnostic test A identifies competent learners who then do the enrichment activities. It also identifies students who have not reached mastery level, diagnoses their learning difficulties and prescribes specific corrective activities. Corrective activities include revision of text, study of alternative materials, additional exercises and peer/teacher tutoring to ensure students master the units when diagnostic test B is applied.

To return to Carroll's learning factors, the procedures of diagnostic tests and corrective activities directly improve the quality of instruction, factor (4), and also allow students the opportunity and time to learn the task, factor (2). Thus, the procedures indirectly heighten and ensure students the ability to understand the instruction, factor (5), and consequently, students are motivated to perseverance to master the instruction, factor (1), (willing to spend the time to learn).

It is argued that the mastery learning strategy heightens achievement, increases learning rate and self-esteem and improves attitudes (Glathorn 1987). Application of the mastery learning strategy in learning biology in a secondary school in Hong Kong has produced higher student achievement (Chung 1979). This strategy should be particularly helpful in learning map work which requires abstract thinking and mathematical calculation and the successful learning of subsequent units depends on mastery of preceding ones. The mastery learning strategy is incorporated into the group learning strategy in the teaching of map work in this study.

8.3 The cooperative learning model

The cooperative learning model is used in the traditional classroom of one teacher teaching many students. Its special features are that students are organized into groups of 4 or 5 students of different abilities and students are rewarded on the basis of individual and group work. This encourages cooperation and mutual help among group members. It is unlike the traditional classroom which encourages competition where someone wins and everyone else loses; or the modern individual education plan (developed particularly to help the slow students) in which a student follows his/her own study plan, with no competition and no interdependence, i.e., no incentive to work against or with anyone else (Reinharts 1984).

The cooperative learning model has two important elements to make it successful: positive interdependence and individual accountability (Slavin 1987). Positive interdependence is achieved through a group task, division of labor, distributed/rotating leadership and a group reward. Individual accountability is achieved by requiring each individual to master the learning material and thereby help to gain the group reward (Dishon & O'Leary 1984; Glathorn 1987; Lockwood 1988).

Cooperative learning improves learning, raises self-esteem and enhances inter-group relations (Johnson & Johnson 1984; Kinney 1989). A review of 27 studies on cooperative learning in grades 7 to 12 indicated an overall success rate of 68% in favor of cooperative over control classes (Newmann & Thompson 1987). Another review of 70 studies on cooperative learning in elementary and secondary schools showed 57% had significantly higher achievement in cooperative than control classes (Slavin 1987).

Cooperative learning has the following advantages (Johnson & Johnson 1984; Ascher 1986; Whitman & Fife 1988):

- (1) students learn better and to a higher and deeper level (conceptual learning and problem-solving) through sustained group effort. Weak students learn because peers are able to explain the material at the right level. Strong students benefit because they gain a better understanding through organizing and explaining the materials to their weaker classmates;

- (2) students are more motivated to achieve higher expectations for success based on collaborative activity;
- (3) students develop more positive attitudes towards the subject area;
- (4) students develop higher self-esteem founded on self-acceptance because they are more successful in school work and are better liked by others;
- (5) students develop closer relationship with classmates regardless of ability level, sex or social class;
- (6) students develop greater interpersonal skills. Cooperative learning encourages and provides for a daily, intense, teacher-sponsored interpersonal contact among diverse students.

The teaching processes of the cooperative learning model are (Foyle et al. 1989):

- (1) specify learning objectives and decide on content;
- (2) determine precise cooperative learning techniques (organization of learning task, assessment and reward);
- (3) assign students of different abilities to groups;
- (4) arrange classroom to facilitate group interaction;
- (5) teach group skills to assure groups run smoothly;
- (6) explain to students the purpose, procedures and activities of learning;
- (7) teach unit 1;
- (8) monitor group learning and provide assistance and clarification as needed;
- (9) evaluate student learning to ensure mastery;
- (10) reward group members for group success.

In this study, the cooperative learning model was adopted in combination with the mastery learning strategy to form the group learning strategy in the teaching of map work. Students in a group cooperated to help one another to master the materials in the learning and corrective activity sessions. To return to Carroll's learning factors, mutual help in the group brings variety of learning situations to improve the quality of instruction, factor (4), and increases the time and opportunity to learn, factor (2); thus group learning heightens students' ability to understand the instruction, factor (5), and consequently, students are motivated to perseverance, factor (1), to master the material.

8.4 The learning team model

Glasser's (1972) learning team model is akin to the cooperative learning model in rationale and operation. The following explanation about the learning team model will confirm further the usefulness of the group learning strategy. Glasser in his book "The Identity Society" explained that shortly after World War II, the newly affluent, television-saturated western world moved rapidly into a society in which the major concern of most people was no longer survival. In the "Identity Society", almost everyone became much more concerned with the psychological needs of love, belonging, fun, freedom and power (Glasser 1986).

Students like everyone else have the above psychological needs. Secondary school students passing through adolescence have stronger needs of love, belonging, fun, freedom and power. Glasser advocates the learning team model in which students work together in small learning-teams of 2 to 5 students made up of low, middle and

high achievers. The learning team model is likely to succeed in motivating almost all students to work for the following reasons (Glasser 1986; Wubbolding 1989): (1) students can gain a sense of love, belonging and fun being in a learning team, (2) students are provided with the initial motivation to work, and as they achieve academic success, those who have not worked previously begin to sense that knowledge is power and then want to work harder, (3) stronger students find it need-fulfilling to help weaker ones because they want the power and friendship that go with a higher-performing team, (4) weaker students find it need-fulfilling to contribute as much as they can to the team effort because now whatever they can contribute helps; when they work alone, a little effort gets them nowhere, (5) students need not depend only on the teacher, they can and are urged to depend a great deal on themselves, individual and team creativity, this gives them both power and freedom, and (6) learning teams can provide the structure for cooperation that will help students to get past the superficiality and to learn enough in depth to make the vital knowledge-is-power connection.

8.5 The group learning strategy

The group learning strategy in teaching of map work is a combination of the mastery learning strategy, the cooperative learning model and the learning team model mentioned in the above sections. Two important features are, firstly, students are divided into groups of 4 consisting of low, average and high achievement helping one another mutually and they will be rewarded for group effort. Secondly, students learn map work in small units, slow students master the material by doing corrective exercises and being helped by other members of the group. The processes are as follows (figure 8.2):

- (1) specify the learning objectives of the map work lessons to students;
- (2) determine precise group learning techniques. In group learning strategy, students sit in groups of 4 in class. They also meet as a group to revise and do homework after class;
- (3) rotate roles of organizer (arranges meeting time and place), checker (checks completion of individual and group work), recorder (keeps record of individual and group work) and praiser (looks for, praises and gives cards to members for improvement or effort) among members to give every member a chance to lead and cooperate;
- (4) arrange classroom to facilitate group interaction. The teacher arranges students to sit in groups in the classroom or geography room. He also books rooms for students for after-school use;
- (5) teach students group skills, such as to listen to others, express themselves and request for help politely;
- (6) explain to students the purpose, procedures and activities of group learning;
- (7) teach unit 1;
- (8) monitor and provide assistance and clarification in group learning in class and in after-school sessions;
- (9) apply diagnostic test A;
- (10) students who do not reach mastery level do corrective exercises being helped by other members. Students who master the material teach weaker peers or do enrichment activities;

- (11) weak students do diagnostic test B to ensure mastery of materials;
- (12) all students do unit test;
- (13) all members are rewarded for group improvement or success;
- (14) teach unit 2;
- (15) repeat processes (8) to (13) for unit 2.

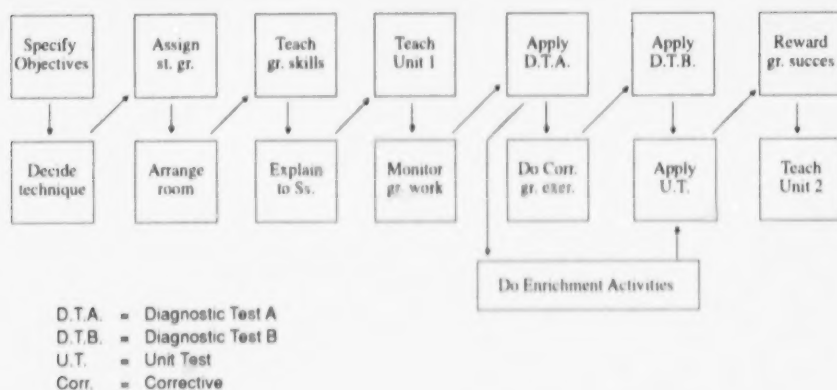


Figure 8.2 Teaching processes of group learning strategy

8.6 The investigation

8.6.1 The sample

Two hundred and twenty-eight students (113 males and 115 females) from a secondary school participated in the investigation. It included two form 1 classes (grade 7), two form 2 classes (grade 8) and two form 3 classes (grade 9). All belonged to high achievement classes as compared to other classes in the school. The classes of each form were taught by the same teacher. One class from each form was assigned at random as the experimental class while the other the control class with the hope that the two groups would be made equivalent.

8.6.2 Method of study

The experimental classes learned map work following the procedures of the group learning strategy described above. Students worked in groups and were encouraged to help one another in addition to getting help from the teacher. They learned map work unit by unit, each of which was followed by exercises, diagnostic test and supplementary exercises to ensure they reach mastery level. The control classes followed traditional classroom procedures and they were just asked to get help from

the teacher when needed. In form 1, three map work units were taught (scale, direction and grid reference). In form 2, two units were taught (contours and contour features). In form 3, map symbols and map interpretation were taught. The whole process took 6 class periods for each form.

Halfway through learning the materials, students of both the experimental and control classes took test 1, and after learning all the map work materials, they took test 2. Analysis of variance with controlling intelligence will indicate any differences between the experimental and control classes which will be due to the effect of the group learning strategy.

8.6.3 Procedure

The investigators went through the procedures of the group learning strategy, namely, specified learning objectives and mastery level, described content, divided and sequenced materials into units, devised appropriate strategies, set diagnostic and other tests, prepared supplementary exercises and devised forms for checker, recorder, and praiser's use.

Before the investigation, students took the Raven's Progressive Matrices (RPM, Short Form 1965) which measures intelligence (reasoning power). The RPM was commonly used in Hong Kong and was found reliable (Siu et al. 1979). Students also completed a sociometric measure both before and after the investigation. Each student was asked to choose 3 classmates with whom he/she would like to form a study group and he/she could also name classmate(s) he/she would prefer not to be in his/her group if that was the case. The result of the sociometric measure, besides for comparison of the experimental and control classes was also used for grouping of students of the experimental classes.

Then the experimental classes learned the map work units following the group learning strategy procedures while the control classes following the traditional classroom procedures. Halfway through the learning, all students took test 1 and at the end of the learning, they took test 2.

8.6.4 Results and discussion

A. Comparison of achievement between the experimental and the control classes

Result of the different forms will be analyzed separately in order to find out whether the group learning strategy has a different effect on learning different topics of map work. A comparison of test scores between the experimental and control classes was made with students' intelligence (RPM scores) controlled (table 8.1).

For form 1, the experimental class did better than the control in test 1 ($F = 2.632$, $df = 1/71$, not significant) and significantly better in test 2 ($F = 3.858$, $df = 1/71$, $p = .053$). For form 2, the control class did significantly better than the experimental in test 1 ($F = 9.817$, $df = 1/77$, $p = .002$), but the experimental class did better than the

control in test 2 ($F = .780$, $df = 1/77$, not significant). For form 3, the experimental class did significantly better than the control in test 1 ($F = 8.155$, $df = 1/71$, $p = .006$) and also significantly better in test 2 ($F = 14.386$, $df = 1/71$, $p = .000$).

Table 8.1 *A comparison of scores between the experimental and control groups*

Form	Test	Mean X Exp. Gr.	Mean X Contr. Gr.	F	df	p
1	1	11.97	10.19	2.63	1/71	.109
	2	13.66	11.78	3.86	1/71	.053*
2	1	22.30	25.53	9.82	1/77	.002*
	2	25.95	24.88	0.78	1/77	.380
3	1	18.00	14.81	8.16	1/71	.006*
	2	19.34	15.94	14.39	1/71	.000*

Thus, the group learning strategy is effective in helping students to learn map work on the whole. For form 1, the experimental class learned better, because it scored higher marks compared to the control class in both tests 1 and 2, with the latter achieving significance. Similarly, for form 3, the experimental class learned better because it scored significantly higher marks in both test 1 and test 2. For form 2, the control class scored significantly higher marks than the experimental class in test 1, however, the experimental class caught up and beat the control class in test 2 though the difference in scores was not significant. A possible explanation for the phenomenon in form 2 is that the map work units of contours and contour features were taught at the beginning of the first term when students did not know one another well enough to cooperate as much for the group learning strategy to take effect more fully. Another possible explanation may be the map work units of contours are too abstract for students to grasp fully to be able to explain them to their peers.

B. Comparison of change in social relationships between the experimental and the control classes

For form 2, the map work units were taught early in the term, so no sociometric measure was done. For forms 1 and 3, an analysis of the sociometric measure indicated that the experimental classes had 8 friendless students at the beginning but 6 of them made friends gradually, leaving only 2 still friendless at the end of the investigation. Whereas the control classes had 7 friendless students at the beginning, only 2 of them made friends leaving 5 still friendless. This indicated that the experimental class students tended to make friends more easily compared to the control class students though there is no measure to indicate significance.

The experimental classes made a total of 15 rejects before the investigation and the number was increased to 17 at the end. The control classes made 18 rejects before the investigation and the number was reduced to 16. This indicated that the group learning strategy which might help students to make friends but it was not likely to get them to reject some classmates less though again there is no measure to indicate significance.

C. Conclusion and recommendation

The group learning strategy which is a combination of the mastery learning strategy and the cooperative learning model is more effective in helping junior secondary students to learn map work compared to the traditional learning method. More students learning through this strategy obtained higher scores in tests than otherwise. They tended to make friends more easily though they did not necessarily reject some of their classmates less. The group learning strategy not only encourages students to master materials in small sequential steps, it also helps to mobilize students to help one another. In this way, peers complement the teacher who cannot give slow students enough timely individual help because of an overload of teaching and administrative duties. Also students tend to make friends more easily, thereby making learning in school more interesting.

The effects of the group learning strategy on learning need to be further researched in Hong Kong, including the writing up and testing of teaching materials, supplementary exercises and diagnostic tests, the development and teaching of appropriate interpersonal skills to students. All these need much time, effort and expertise, but it is worth the effort as there is indication that the group learning strategy will help students to learn more effectively.

References

- ASCHER, C. (1986), Cooperative learning in the urban classroom. New York: Eric/Cue Digest. 30, Eric Clearing House on Urban Education.
- BLOOM, B.S. (1968), Learning for mastery. *Evaluation Comment* 1 (2).
- CARROLL, J.B. (1963), A model of school learning, *Teachers College Record* 64, pp. 723-733.
- CHUNG, Y.P. (1979), The application of mastery learning strategy in the teaching of secondary school biology, *Education Journal. The School of Education, The Chinese University of Hong Kong* 7, pp. 91-101.
- DISHON, D. & P.W. O'LEARY (1984), A guidebook for cooperative learning: A technique for creating more effective schools. Holmes Beach, Florida: Learning Publications.
- FOYLE, H.C., L. LYMAN & M.A. MOREHEAD (1989), Interactive learning: Creating an environment for cooperative learning. Paper presented at the Annual Conference of the Association for Supervision & Curriculum Development. Orlando, Florida.
- GLASSER, W. (1972), *The identity society*. New York: Harper and Row.
- GLASSER, W. (1986), Control theory in the classroom. New York: Harper and Row (Chap. 6).
- GLATHORN, A.A. (1987), Curriculum renewal. Association for Supervision and Curriculum Development. Alexandria, Virginia.

- GUSKEY, T.R. (1985), *Implementing mastery learning*. Belmont, California: Wadsworth.
- HONG KONG EXAMINATIONS AUTHORITY (1980-89), *The Hong Kong certificate of secondary education examination annual reports*. Hong Kong: The Hong Kong Examinations Authority.
- JOHNSON, D.W. & R.T. JOHNSON (1984), *Cooperative small group learning*. National Association of Secondary School Principals. Reston, Virginia.
- KINNEY, J.H. (1989), *A study on the effects of cooperative learning program on achievement of ninth-grade multi-cultural biology classes*. Paper to the Alexandria City, Virginia School Board regarding teaching techniques aimed at improving existing programs.
- LOCKWOOD, A.T. (1988), *Cooperative learning*. National Centre of Effective Secondary Schools, Resource Bulletin 4.
- NEWMANN, F.M. & J.A. THOMPSON (1987), *Effects of cooperative learning on achievement in secondary schools, a summary of research*. Madison: Wisconsin Center of Education Research.
- RAVEN, J.C. (1965), *Advanced progressive matrices sets I and II*. London: Lewis.
- REINHARTS, J. (1984), *Principles on effective teaching and the cooperative classroom. Analysis and action series*. Washington, D.C.: National Educational Association.
- SIU, P.K. et al. (1979), *The effects of the medium of instruction on student cognitive development and academic achievement*. School of Education, The Chinese University of Hong Kong.
- SLAVIN, R.E. (1987), *Cooperative learning student teams, What research says to the teacher* (2nd Ed.). Washington, D.C.: National Education Association.
- STIMPSON P. & Y.L. KWAN (1990), *Teaching geography in and about Asian Pacific countries. Abstracts and papers of the Commission on Geographical Education of the International Geographical Union*. Hong Kong: University of Hong Kong/IGU.
- WHITMAN N.A. & J.D. FIFE (1988), *Peer teaching: To teach is to learn twice*. ASHE-ERIC Higher Education Report, No. 4.
- WUBBOLDING, R.E. (1989), *Using reality therapy*. New York: Harper and Row.

9 EFFECTIVENESS OF GAMING METHOD IN TEACHING GRID REFERENCES AND CONVENTIONAL SIGNS: AN EXPERIMENTAL STUDY

Yee-Wang Fung

9.1 Introduction

The gradual shift in the emphases of geographical education which started in the mid-sixties has led to innovative changes in teaching strategies. One of the innovations was the introduction of simulations into the geography room.

It has been pointed out by Taylor & Walford (1978) that simulation techniques are centuries old, but their importation into general education is only a recent phenomenon. The more overt and conscious use of simulation techniques in the field of research and education was probably developed in the 1930s, when there was a growing interest in small-group behaviour, by psychologists, psychiatrists and sociologists. Thus McAleese (1978, p. 11) says, "Although the power of play in education has long been acknowledged, and indeed may be traced back via Bruner, Piaget, Dewey and Rousseau to Plato, it is only in recent years that serious thought has been given to the practical problems and possibilities inherent in its application".

According to Taylor and Walford (1978), simulation covers role-play, gaming and machine simulation. Taylor (1977) further offers a classification of simulation techniques based on a reality-abstraction scale. Six techniques are identified which range from case study on the reality end through in-basket method, incident process, role playing and gaming simulation to machine simulation on the abstraction end. Jones (1985), however, argues that there are clear distinctions between simulations, games and role plays. Seidner (1978) suggests that simulations are abstractions and simplifications of the real world, focusing upon particular aspects of the referent system rather than upon all of its elements; play is a free, make-believe activity; and a game is an activity with a set of rules which limits the means whereby the objectives agreed upon by the participants may be attained.

There is ample evidence to substantiate the claim that games provide a superb means of motivating children to participate actively in the process of learning, but they have not been found to be particularly effective in imparting knowledge, teaching problem-solving skills and inducing critical thinking (Seidner 1978; Cherryholmes 1966).

In summarizing the studies in this field Seidner reported that the delayed retest scores from several studies indicated that students retained more of what they learned from participation in the simulation than did students taught by traditional methods. Moreover, as there is a curvilinear relationship between tension and learning, he believed that gaming situations would "provide just enough tension in the form of wanting to win, without invoking excessive worry about failure, as in a testing situation" (Seidner 1978, p. 20).

Higher motivation and better performance are achieved not only in individual competitions, but also in group competitions (Julian & Perry 1967). Walford (1969, p. 29) suggests a new dimension for consideration when he says, "Some of the elements of game playing are of considerable importance in developing relationships and personalities in the classroom. Games in which groups of pupils need to discuss and make decisions are practical and simple examples of the kind of democratic procedures which Dewey thought important".

Notwithstanding what was said above, both enthusiasts and critics have cautioned the teaching profession that not only may games be overused but they may also provide a disservice to education. As Walford (1969, p. 30) pointed out, "An invaried diet of games or an over-extension of them beyond a pupils' span of concentration is as unwise as a constant diet of note-taking or textbook reading". While Kasperson (1968) was concerned about the danger of winning rather than learning being emphasized in a gaming situation, Lundy (1985) noted that the competitive nature of the game might cause the students to behave in such a way that poor team positions and poor understanding of the underlying principles of the game would result. Moreover, Radley (1979) argued that not all participants in simulations and games benefited from the experience and that some might even be psychologically damaged by being involved in this mode of tuition because of the group pressure present in team-based simulation games.

Finally it may be noted that the advantages and disadvantages of simulations and/or simulation games as a teaching strategy in general have been neatly summarized by Taylor & Walford (1978) and as a teaching strategy for geography, by Spicer (1973).

The present study aims to ascertain the effectiveness of three different teaching modes, namely, group competition, individual competition and traditional classroom instruction, in teaching a geography lesson on grid references and conventional signs to S1 (secondary 1, equivalent to grade 7) students.

9.2 Methodology

Three typical secondary schools in Hong Kong were invited to participate in the study, involving a total of nine classes with about forty students in each class. Their academic standard ranged from slightly below average to slightly above average. The classes chosen in each school also belonged to the average category within the school in terms of academic standard.

The three classes in each school were randomly assigned the roles of experimental group 1 (group competition), experimental group 2 (individual competition) and control group (traditional instruction). All the students were asked to answer a questionnaire in the second week of September, which yielded two indices: (1) democratic tendency; (2) geographical interest. In the third week of September the lesson was taught to all the groups in a random order, using one double period (70-80 minutes) for experimental groups and two single periods (35-40 minutes each) or one double period for the control group. Test 1 (see appendix 3) was administered to the students in the first geography period after the lesson. Immediately after the test, the answers were given and explained to them. For the experimental groups, the competition results were also announced. Two weeks later, when test 2 (see appendix 4) was administered, the students were also asked to repeat the questionnaire they had answered four weeks ago. The procedures outlined above in the end produced the following scores for each group:

T1	(learning outcome	2-3 days after the lesson)
T2	(learning outcome	2-3 weeks after the lesson)
D1	(democratic tendency	1 week before the lesson)
D2	(democratic tendency	3 weeks after the lesson)
G1	(geographical interest	1 week before the lesson)
G2	(geographical interest	3 weeks after the lesson)

In the questionnaire (see appendix 5), the students were asked to respond to three different situations, for each of which they were given three responses (autocratic, democratic and laissez-faire) to choose from. The number of responses showing a democratic tendency chosen by the students in each group was summed and expressed as a percentage of the total number of responses.

The geographical interest index was obtained by combining the scores of two questions: the first one asked the students to rate their own interest in geography on a 5-point scale; the second one asked them to rank geography together with eight other school subjects. The two scores were combined in the ratio of 2:1.

The lesson for the experimental groups was built around a game on grid references and conventional signs called "Protect Your District" (see appendix 2). The students were given the role of a security officer who knew that a gang from outside would come and create trouble in his district. As a security officer, he had to decide where he should station his squads. A map of the district was given, and he had to locate on the map the five most important strategic points, give their 6-digit grid references and name the conventional signs. The game was played in groups for experimental group 1 and individually for experimental group 2.

Teaching materials included a conventional textbook explanation of grid references, a student game instruction sheet, a large-scale wall-map with grids and conventional signs and a small-scale wall-map with longitudes and latitudes. Teaching strategies included explaining, drilling, reading, gaming and discussing. The lesson was designed by the author, using the game as the centrepiece of this unit of work (Taylor & Walford 1978). In designing the lesson, the following advice by Seidner (1978, p. 23) was borne in mind, "It would appear that the best way to insure learning from simulation games would be to design the game in such a way as to make the desired learning requisite to playing the game.....".

A trial lesson was conducted by the author which was video-taped for showing to the teachers who helped to teach the experimental classes. The lesson plan adopted for use in the present study was slightly different from the original one which was described in some detail elsewhere (Fung 1988). The revised plan consisted of the following five major steps (see appendix 1):

1. *Motivation* (5 mins.): arousing students' interest in maps.
2. *Development* (15 mins.): teaching the students conventional signs and grid references.
3. *Consolidation* (10 mins.): drilling the students.
4. *Application* (20 mins.): running a gaming session.
5. *Conclusion* (15 mins.): discussing the game and revising the lesson.

As for the control group, the teachers were asked to teach the classes as they normally did. In this case, the traditional approach would involve the following five steps:

1. *Motivation*
2. *Development* (teaching the students)
3. *Application* (students doing exercise)
4. *Consolidation* (discussing the exercise and setting homework)
5. *Conclusion*

The results were analysed and used to test the null hypotheses that no significant differences existed:

1. between the T1 scores of the three groups;
2. between the T1 and T2 scores of each group;
3. between the T2 scores of the three groups;
4. between the D1 scores of the three groups;
5. between the D1 and D2 scores of each group;
6. between the D2 scores of the three groups;
7. between the G1 scores of the three groups;
8. between the G1 and G2 scores of each group;
9. between the G2 scores of the three groups.

9.3 Result

Table 9.1 shows that in test 1, the experimental group 2 (individual competition) scored 7.235 out of 10, followed by the experimental group 1 (group competition) with a score of 6.375 and the control group (traditional instruction) with a score of 5.451. The differences between these statistics were found to be significant at the 0.01 level of confidence when the F test was applied. Thus the null hypothesis that there were no significant differences between the T1 scores of the three groups was refuted. Further analyses show that the difference in T1 scores between the experimental group 1 and the experimental group 2 (0.860) and that between the experimental group 1 and the control group (0.924) were significant at the 0.05 level of confidence, while that between the experimental group 2 and the control group (1.784) was significant at the 0.01 level of confidence.

Table 9.1 Learning outcomes of two experimental groups and one control group

	Experimental group 1 (group competition)	Experimental group 2 (individual competition)	Control group (traditional instruction)	F-test
<i>Test 1</i>				
- N	112	119	113	F = 10.430**
- Mean score	6.375	7.235	5.451	df=2/341
- SD	3.250	2.782	2.882	
<i>Test 2</i>				
- N	112	119	113	F = 0.040
- Mean score	6.223	6.126	6.142	df=2/341
- SD	2.967	2.451	2.955	
<i>t-test</i>				
- t	0.365	3.263**	-1.780	
- df	222	236	224	

The differences between T1 and T2 scores for the experimental group 1, the experimental group 2 and the control group were respectively 0.152, 1.109 and -0.691. When the t-test was applied, only the difference for the experimental group 2 (1.109) was found to be statistically significant at the 0.01 level of confidence. Therefore, only for this group the null hypothesis that there were no significant differences between the T1 and T2 scores of each group was refuted.

As for the T2 scores, because the range was small (from 6.126 to 6.223), none of the differences between the three groups were statistically significant. The null hypothesis that there were no significant differences between the T2 scores of the three groups was therefore retained.

Likewise, little differences were found among and between the D1 and D2 scores for the three groups of students. These scores ranged from 84.17% to 90.56%. The same was also true of the G1 and G2 scores, which ranged from 4.504 to 4.903 out of a total 9.500. Thus all the six null hypotheses from no.4 to no.9 were retained.

9.4 Discussion

The T1 scores show that the students in the experimental groups learned significantly better than those in the control group. Between the two experimental groups, the one using an individual competition approach achieved a significantly higher score than the one adopting a group competition method. Since these differences were significant at the 0.05 or a higher level of confidence and there was no reason to believe that there were already differences between the samples before the lesson was taught, it would be logical to conclude that different teaching strategies did produce different learning outcomes.

However, these learning outcomes were only temporary. What was learned might

be forgotten after a period of time. How much of what was learned was retained might also be related to how it was learned. The present results show that for the control group, not only was there no decrease in the test score after about two weeks, but there was in fact some increase, which was, however, not found to be statistically significant. There was a negligible decrease for the experimental group 1, but the decrease for the experimental group 2 was quite substantial which was found to be significant at the 0.01 level of confidence. The conclusion that can be drawn at this point is that although the students learned better with the individual competition approach, they also forgot more of what they had learned within a period of about two weeks.

It is interesting to note that the T2 scores obtained by the three groups were very much the same, ranging only from 6.126 to 6.223. The differences were obviously statistically insignificant. Enthusiasts will undoubtedly be disheartened by this result which shows that although the gaming approach produced better learning outcomes at first, its long term effects were less convincing.

The results obtained in this experimental study seemed to suggest that no matter what teaching strategy was employed, the final results would not differ too much. Feedbacks from some of the teachers involved in teaching the classes showed that students in the experimental groups generally found the lesson exciting and interesting, but they had not been given any homework to do after the lesson as their fellow students in the control group.

Thus the following explanation may be advanced to account for the present phenomena. The three different teaching strategies did produce different learning outcomes, but because there were no follow-up activities for the experimental groups, they forgot a part of what they had learned after a while. That the students taught by the individual competition approach forgot a significant amount of what they had acquired whereas those taught by the group competition approach forgot only a negligible amount may be attributable to the fact that the former had not gone through a group discussion process as the latter. Group discussion helps students to consolidate what they have learned. Although the difference in the T1 and T2 scores of the Control Group was not statistically significant, it suggested that some follow-up activities after the lesson might help to consolidate or even improve the learning taken place in the classroom.

All the null hypotheses on democratic tendency and geographical interest were retained mainly because a double-period lesson alone was insufficient to cause any noticeable changes in attitude. Attitude change is a long term process. It is therefore suggested that in order to ascertain the effects of the gaming method in regard to attitude changes, sustained longitudinal studies have to be conducted (Walford 1985; Twelker 1977). Moreover, the art of evaluation also needs to be improved (Percival 1978).

9.5 Conclusion

Findings in this experimental study show that the gaming approach did produce desirable short term cognitive learning outcomes, but the differences in achievement between the three groups, two experimental and one control, were completely

ironed out within two to three weeks. It is believed that the lack of group discussions and/or follow-up activities was the main cause of the slackening of the experimental groups.

Mainly because the present study is an isolated study, the affective effects of the gaming method were not observable. The mounting of longitudinal studies will no doubt go a long way to help solve this problem.

It is apparent that the sharpening of research techniques alone will not completely solve the problem. There must also be teachers who are able to integrate gaming activities effectively into classroom teaching. Furthermore, there must be a good supply of geographical games which can be used effectively in the classroom.

References

- CHERRYHOLMES, C.H. (1966), Some current research on the effectiveness of educational simulations. *American Behavioural Scientist* 10 (2), pp. 4-7.
- FUNG, Y.W. (1988), Civic education through geography: A case-study. In: R. Gerber & J. Lidstone, eds., *Developing skills in geographical education*. Brisbane: Jacaranda-Wiley.
- JONES, K. (1985), *Designing your own simulations*. London: Methuen.
- JULIAN, J.W. & F.A. PERRY (1967), Cooperation contrasted with intra-group and inter-group competition. *Sociometry* 30 (March), pp. 79-90.
- LUNDY, J. (1985), The effects of competition in business games. In: M.V. Ments & K. Hearnden, eds., *Effective use of games and simulations*. Leicestershire: Loughborough University of Technology.
- MCALLEESE, R. (1978), Introduction. In: R. Mcalleese, ed., *Perspectives on academic gaming and simulation 3*. London: Kogan Page.
- PERCIVAL, F. (1978), Evaluation procedures for simulation/gaming exercises. In: R. Mcalleese, ed., *Perspectives on academic gaming and simulation 3*. London: Kogan Page.
- RODLEY, G. (1979), The dynamics of groups in gaming. In: J. Megarry, ed., *Perspectives on academic gaming and simulation 4*. London: Kogan Page.
- SEIDNER, C.J. (1978), Teaching with simulation and games. In: R.L. Dukes & C.J. Seidner, eds., *Learning with simulations and games*. Beverly Hills: Sage.
- SPICER, B.J. (1973), The use of simulation games in secondary school geography. In: D.S. Biddle & C.E. Deer, eds., *Readings in geographical education*. Australia: Whitcombe and Tombs.
- TAYLOR, J. (1977), Instructional gaming procedures in planning education. In: J. Megarry, ed., *Aspects of simulation and gaming*. London: Kogan Page.
- TAYLOR, J. & R. WALFORD (1978), *Learning and the simulation game*. Milton Keynes: The Open University.
- TWELKER, P. (1977), Some reflections on the innovation of simulation and gaming. In: J. Megarry, ed., *Aspects of simulation and gaming*. London: Kogan Page.
- WALFORD, R. (1969), *Games in geography*. London: Longman.
- WALFORD, R. (1985), Simulation: after the honeymoon - what next? In: M.V. Ments & K. Hearnden, eds., *Effective use of games and simulations*. Leicestershire: Loughborough University of Technology.

Appendix 1

Lesson plan for playing the "Protect Your District" game

OBJECTIVE AND CONTENT	TIME (min)	TEACHER ACTIVITY	PUPIL ACTIVITY	REMARK
Motivation				
1 To arouse pupils' interest in maps and awareness of the need to master the skills of reading maps	5	Talking; questioning	Listening; answering	Divide the pupils into groups of five each
Development: (a) conventional signs				
2 To help pupils understand that conventional signs are map-readers' language	10	Talking; questioning	Listening; answering	
3 To enable pupils to learn some basic conventional signs				
Development: (b) grid reference				
4 To help pupils to understand the function of grid reference	10			
5 To enable pupils to use 4-digit grid reference		Talking; questioning	Listening; answering	
6 To enable pupils to use 6-digit grid reference				
Consolidation				
7 To help pupils consolidate their knowledge of conventional signs and grid reference through (a) blackboard exercise and (b) silent reading	15	Questioning; observing	Answering; reading	Distribute reading material
Application				
8 To stimulate pupils' interest in learning and to enhance their learning effects through a competitive game	15	Talking; observing	Listening; playing	Pupils to put away reading material; teacher to distribute game's sheets (white for personal use; green for scoring); pupils to hand in the green sheet once they have finished
Conclusion				
9 To help pupils clarify any doubts on conventional signs and grid reference by going over their answers	10	Talking	Listening	

Appendix 2

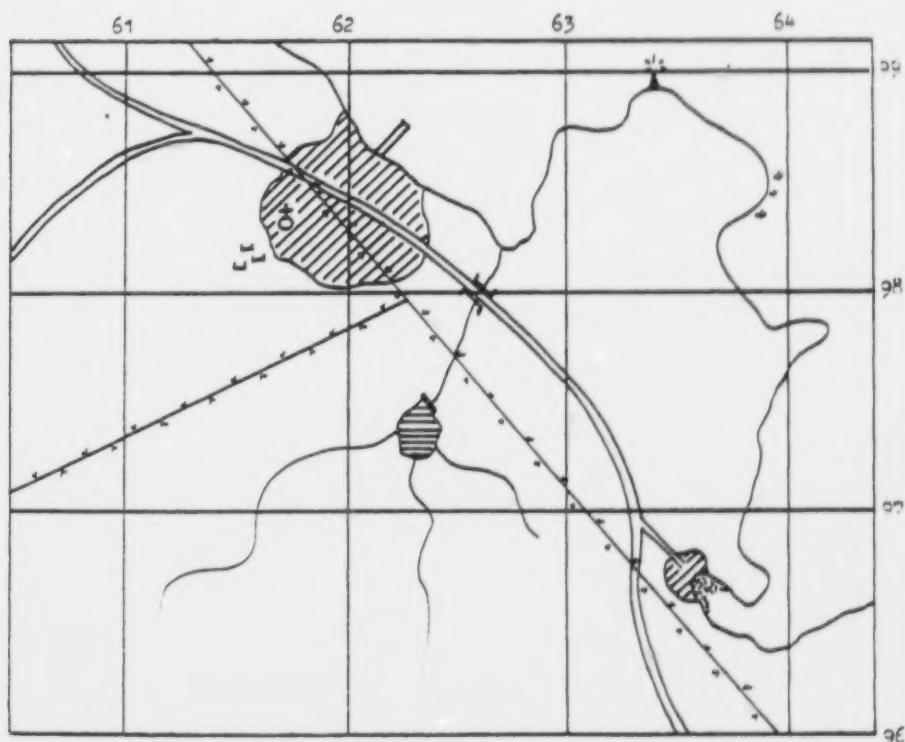
The game "Protect Your District"

Rules

1. The Security Officer knows that a gang from outside will come and create trouble in this district. He has only five squads. If you were the officer, where would you station te squads?
2. There are altogether 3 rounds: Round 1, locate Stations A-E on the map; Round 2, write the grid references of A-E in the table; Round 3, name the conventional signs for A-E in the table.
3. For each correct answer, there is one mark. The highest possible score is 15, and the lowest, 0. If the socres are equal, the time counts.

Code	Location (grid reference)	Object for protection
A		
B		
C		
D		
E		

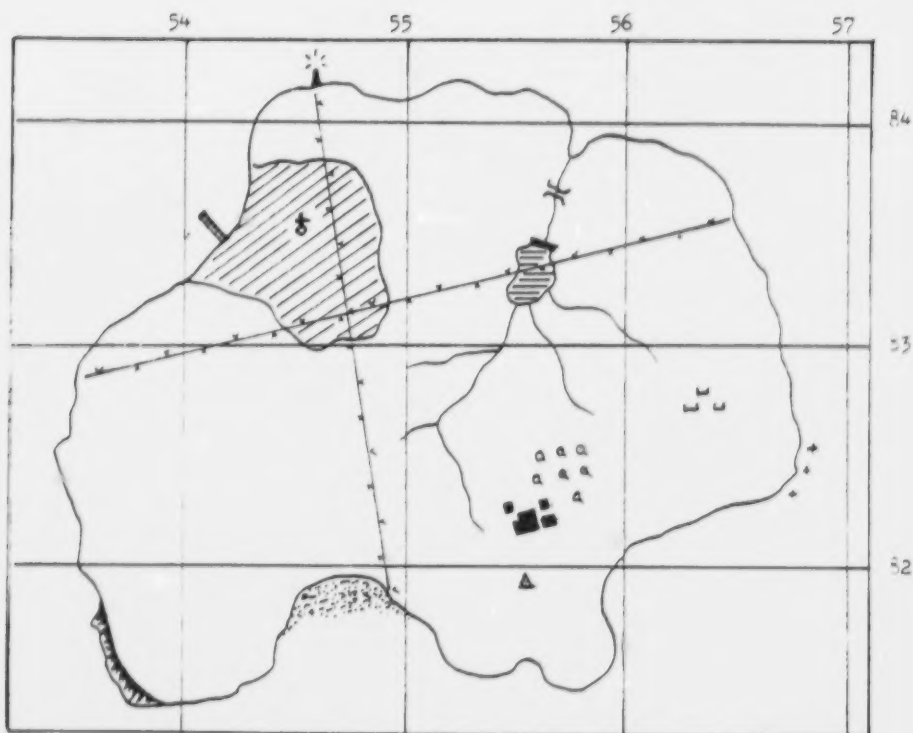
Time: ____ min. ____ sec.



Appendix 3

TEST 1

Name: _____



1. Give the 4-digit reference for the following features:

- (a) Reservoir
- (b) Beach
- (c) Cemetery
- (d) Awash and submerged rocks
- (e) Pier

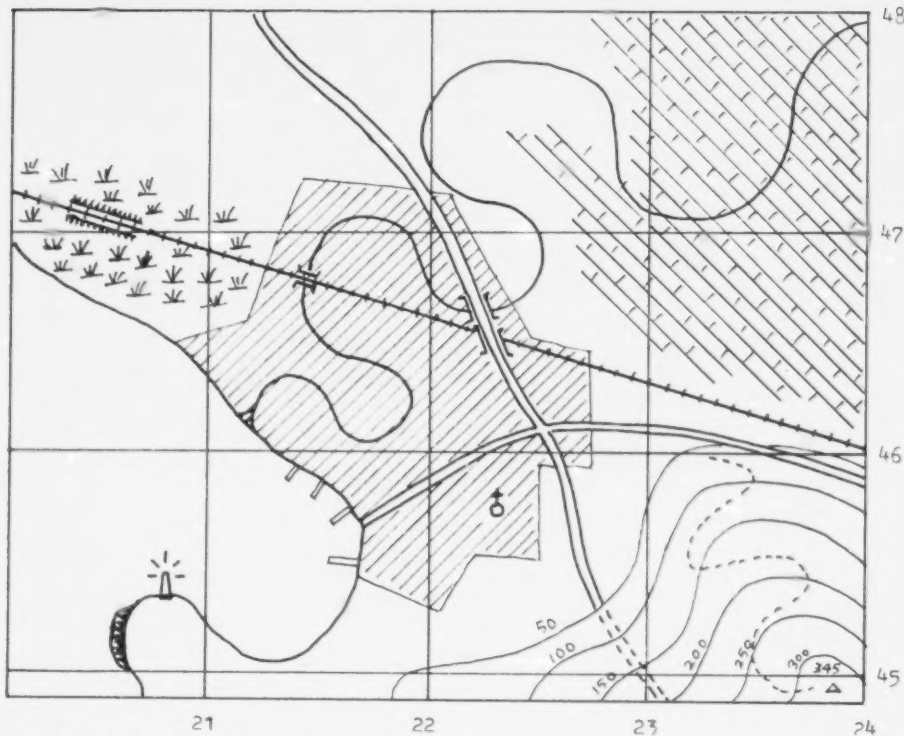
2. Give the 6-digit reference for the following features:

- (a) Bridge
- (b) Church
- (c) Lighthouse
- (d) Where the river enters the sea
- (e) Where the two power lines intersect

Appendix 4

TEST 2

Name: _____



1. What do the following signs represent?

(a) + + + + + + + +

(b) = = = = = = = =

(c) ↓ ↓ ↓ ↓ ↓ ↓

(d) \ / \ / \ / \ / \ /

(e) - - - - - - - -

2. Give the 6 digit reference for the following features:

(a) Church

(b) Cliff

(c) Lighthouse

(d) Railway bridge

(e) Road junction

Appendix 5

QUESTIONNAIRE

1. In a class meeting, after a lengthy discussion the members are still unable to decide on the venue for an outing. As a member of the group, what would you suggest?
 - ___ Let the class teacher/captain decide.
 - ___ Let the members divide into small groups and decide for themselves.
 - ___ Let the members vote on the proposals under the guidance of the class teacher/captain.

2. You are the leader of a small group. When a member chooses to attend to his own business irrespective of what is going on in a group function, what would you do?
 - ___ Let him carry on.
 - ___ Try to find out why he behaves that way.
 - ___ Stop him from attending to his own business.

3. In a meeting, the Chairman expresses his personal view on an issue and requests the members to accept it. If you hold a different view, what would you do?
 - ___ Put forth your own view.
 - ___ Accept his with reluctance.
 - ___ Remain silent, but be prepared to act independently.

4. Are you interested in geography?

1 2 3 4 5

Interested Not interested.

5. Please rank the following subjects in order of your interest in them: "1" for the most interested, "2" for the second most interested, "3" for the third most interested, and so on. Please do not rank those subjects not offered in your school.

___	Chinese language
___	English language
___	Mathematics
___	Integrated Science
___	Chinese History
___	World History
___	Geography
___	Economics and Public Affairs
___	Social Studies

6. Name _____
7. Sex _____

10 ASSESSMENT IN GEOGRAPHY: AN EVALUATION OF THE SOLO TAXONOMY

Philip Stimpson

10.1 Introduction

Geography teachers, as with workers from other discipline areas, expend much effort in assessment. Teachers are expected to be able to design questions which will challenge and interest pupils. At the same time questions should permit the teacher to differentiate between levels of understanding in order to offer remedial help and decide new directions of study for the individual pupil. Just as it is accepted that effective teaching depends on matching content with pupils level of understanding which implies the existence an underpinning in learning theory, so this is also true of assessment. The construction of suitable tests necessitates that we critically review how we assess at present and that we examine new approaches. Innovation in education, however, is littered with seemingly good ideas which have foundered when they have reached the classroom. The research and development element in curriculum change is frequently weak. Insufficient attention is often given to evaluating new ideas before they are introduced and to the methods by which such evaluation can be carried out. This paper focuses on the need for a theoretical base for testing and on the evaluation of the SOLO taxonomy (Biggs & Collis 1982) as an approach to structuring criterion referenced questions in geography.

10.2 The need for theory

The need for theory becomes apparent when many existing questions are examined. The multiple choice question shown in figure 10.1 is typical of many given to Certificate Level (16+) pupils in Hong Kong. What do correct or incorrect answers tell us about our pupils?

In this question, pupils who make the correct choice exhibit the ability to bring two pieces of information together, namely low birth rate and high gross national product (GNP). Similar thinking leads to the rejection of response C. In B and D, however, pupils must integrate the information which is intellectually more demanding. In this multiple choice question, it probably was not serious that some options were inherently more difficult to handle than others because the correct answer was derived through the easier task involved in response A. It would have been of concern if the questions had read: "A -the birth rate is high in countries with a low GNP", and "D - the birth rate of a country decreases with increasing GNP" with B and C as stated. D, with the more difficult thinking task, is the correct answer. There may have been pupils who could have picked out a number of individual pieces of

information but who lacked the sophistication to bring them together; such pupils would fail to demonstrate any ability in terms of the test. The cognitive demand of the subitem is important.

"The diagram shows the birth rate against the Gross National Product (GNP) of 2 countries. What is the association between GNP and crude birth rate?"

- A. Crude birth rate is low in countries with high GNP.
- B. Crude birth rate bears no relation with GNP.
- C. Crude birth rate is high in countries with high GNP.
- D. Crude birth rate of a country decreases as GNP falls."

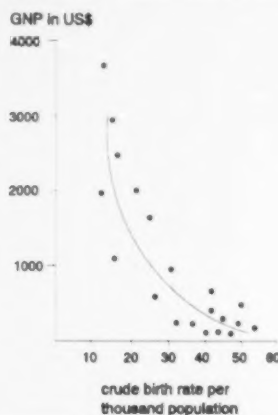


Figure 10.1 A typical multiple choice geography question

Most teachers would now not accept that an incorrect response indicates a pupil has no understanding of the topic. Pupils may have been able to demonstrate some learning in the particular area but at a lower level of difficulty. All attempts to answer a question should be regarded as an achievement and answers need to be classified by level of response rather than simply right or wrong. Tests are needed which differentiate and not just discriminate between pupils. Hence differentiated assessment frequently uses tests which contain graded components. These tests seek to measure responses, not just at one level, but at various levels.

Traditional non-graded tests often discriminate in a negative sense and are ineffective and even possibly counterproductive for a large proportion of the school population. These tests are often set so that about one-third of candidates fail; discrimination is achieved by setters through searching out an appropriate level of obscurity. Candidates who gain high marks often do so by reproducing a lot of knowledge whereas those who produce few facts do less well. The effectiveness of the assessment and what can be interpreted from it is brought into question. In an effective graded test, criteria define what each grade or level signifies. By adopting criterion referenced tests which are appropriately targeted across the

ability range, most pupils achieve something although some achieve more than others. This is patently needed in many education systems where there is a comprehensive scheme of schooling across all abilities. For example in Hong Kong, Certificate Level (16+) geography is taken by 90% of pupils aged generally between 16 and 17 years. The norm referenced expectation is that about two-thirds will succeed and hence the target population is located between the 1st and approximately the 60th percentiles. Within this range, however, there is considerable variation in ability. In Australia, Biddle (1978) estimated that about 35% of 16 year olds probably reveal formal operational behaviour. As many again (37%), however, function at a dominantly concrete level of thinking, many pupils being only able to apply logical thought to problems concerning the real things they experience or have experienced. Graded tests are thus vital if the able are to be encouraged and at the same time the less able are not to be discouraged.

10.3 The SOLO model

In order to place such tests on a sound footing a theoretical base is required. There has to be a graded structure underpinning the tests which reflects the different levels of understanding children acquire. A model is needed which describes increasing competence. Rhys (1972) was amongst a number of people in the 1970s who were interested in Piagetian stages of development and hinted at the problem. It was Bennetts (1981), however, who focused attention explicitly on the issue of progression in the U.K. geography curriculum. Nevertheless, the desire to implement such changes at that time largely preceded the search for underlying principles to structure tests. A theoretical basis can either be derived by the topic-specific technique of phenomenography (Marton 1981; Ramsden 1988), whereby students' responses to a given topic are ordered into categories of differential expertise, or by imposing a general model based on the structural similarities previously established across topics or subject areas. The second approach, as long as it can be shown to apply to the topic addressed in each item, has the advantage that one can speak in a common language across topics and year levels.

Biggs & Collis (1982) attempted to show that students learn quite diverse material in stages of ascending structural complexity that display a similar sequence across tasks. Out of this they proposed a general taxonomy of learning called the "SOLO Taxonomy", SOLO standing for **Structure of the Observed Learning Outcome**. It involves a hierarchy of thinking skills which are differentiated by nature of cognitive ability demonstrated rather than directly the difficulty of the task. The taxonomy makes it possible, in the course of learning a subject, to identify in broad terms the stage at which a student is currently operating.

The following stages are recognised in the model. The ability to:

- make preliminary preparations, but the task itself is not attacked in an appropriate way (prestructural);
- extract one aspect of the problem (unistructural), followed by a situation where several aspects of the task are picked up serially, but are not interrelated (multistructural);
- integrate several aspects into a coherent whole (relational);
- generalise that coherent whole to a higher level of abstraction (extended abstract).

The taxonomy can be used either to classify responses to open-end questions, or to structure objective-type items in the ordered-outcome format (Collis & Davey 1986; Wilson 1989). The latter format is the one illustrated here. In constructing such items, the aim is to ask a series of questions about data contained in a stem in such a way that satisfactory responses require a more and more sophisticated use of the information given in the stem. The following criteria are used to design subitems addressing the stem topic at each SOLO level:

1. *Unistructural*: Contains one obvious piece of information coming directly from the stem.
2. *Multistructural*: Requires using two or more discrete and separate pieces of information contained in the stem.
3. *Relational*: Uses two or more pieces of information each directly related to an integrated understanding of the information in the stem.
4. *Extended Abstract*: Requires use of an abstract general principle or hypothesis which can be derived from, or suggested by, the information in the stem.

The unstructured accumulation of knowledge, of which rote learning would be one (but not the only) example, is emphasised in responding up to multistructural level; structuring knowledge with increasingly abstract frameworks is emphasised through relational to extended abstract responding. The SOLO model has received some interest from geographers already (Biddle 1978; Hammond 1983; Wilson 1988) although only Courtney (1989) has attempted any field trial. This research was, however, relatively small scale involving only 68 young people between the ages of 13 and 20 years.

10.4 SOLO based ordered outcome questions in geography

Geography is concerned with enabling people to make sense of the world around them. Above all it seeks to teach children how to learn about their surroundings. This focuses attention on the fundamental geographical questions of *What, Where, Why, How, What Will* and *How Ought*. These seek to develop powers of description (*What? Where?*), explanation and interpretation (*Why? How?*), prediction (*What will?*) and evaluation (*How ought?*). Description requires pupils to be able to observe and extract either single or linked pieces of information whilst explanation is dependent upon the understanding of relationships which link factors and process elements to the landscape features they determine. Prediction requires that the pupil is able to generalise outside the confines of the landscape relationships which are observed. There is a clear parallel with the SOLO framework.

A second feature of contemporary geography teaching is the importance given to data based learning and problem solving. The hope is that pupils, when confronted with a new question such as why are there factories rather than homes located in a particular area, are able to examine the situation, make sense of the information they have extracted and generate a reasoned interpretation.

SOLO type ordered outcome questions offer the potential of test items which are consistent with this way of thinking in geography within a problem solving frame-

work. They contain a data based stem and a set of four sub-questions, one at each of the four SOLO levels. The stem provides pupils with opportunities to practice data handling skills and the ordered questions are compatible with the *What, Where, Why, How Ought* focus of geography.

10.4.1 An example of an ordered outcome question

The question in figure 10.2 illustrates the general form of such questions using the example of the assessment of learning about the relationship between temperature and latitude. The stem contains two climate graphs and a map showing the location of their respective climate stations. An initial low level task requires pupils to extract one piece information from the data stem; pupils are asked to find the mean annual temperature of Singapore which is marked on one of the graphs. This is a *unistructural* task (U) in SOLO terms. At the next, *multistructural* (M), level pupils are required to state how many times greater is the mean annual temperature of Singapore than Urumqi; in this it is necessary to extract two discrete pieces of information from the stem. In geographical terms, these two levels largely encompass the fundamental geographical questions of *WHAT* and *WHERE*. They are mainly descriptive and are the usual first steps in an inductive approach to the teaching of geography which is frequently advocated today.

At a higher level, as illustrated in subitem (R), pupils are asked to show their understanding, from the data, as to *HOW* and *WHY* average annual temperature changes with latitude. The pupils need to use three pieces of information each of which is directly linked in an integrated understanding of a concept implicit in the stem. The key is the integration of the available data and hence pupils are required to function at a *relational level* (R) of thinking. The ability to generalize can be extended to a higher level. In the fourth subitem (E) pupils are asked to think of other factors which might influence the difference in temperature and hence to place what is presented in the stem in a general factor system. It could be argued, however, that, if the context is familiar, then this answer as at the previous level might be recalled information at a multistructural level. Alternatively, therefore, one might ask pupils at the *extended abstract* (E) level to predict, from the relationship in subitem (R), the temperature of some other location. This requires the pupil to work in an extended abstract mode of thinking.

Ordered outcome questions, therefore, consist of set of data contained in a stem about which four questions are asked. These four questions are arranged in increasing order of difficulty according to the four SOLO levels. The nature of the stem and the four questions are now looked at in more detail.

10.4.2 The stem

Study the two graphs and the map (on page 162)

- (U) From the graph, what is the average annual temperature in Singapore? (Do not calculate it.)
- (M) How many times greater is the average annual temperature in Singapore than Urumqi?
- (R) From the data, what is the relationship between latitude and average annual temperature?
- (E) Is latitude the only factor which is likely to explain differences in temperature? Explain your answer.

Average temperature and rainfall in Singapore (A) and Urumqi (B)

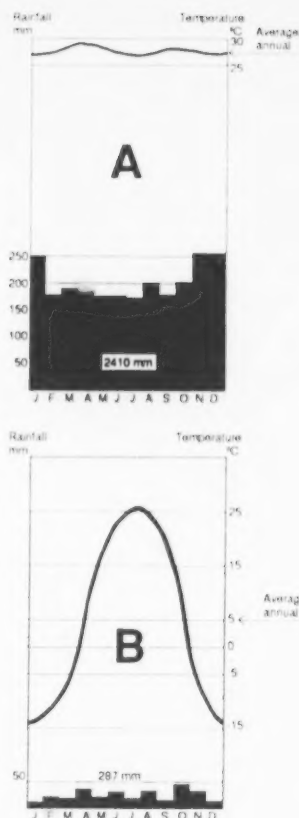
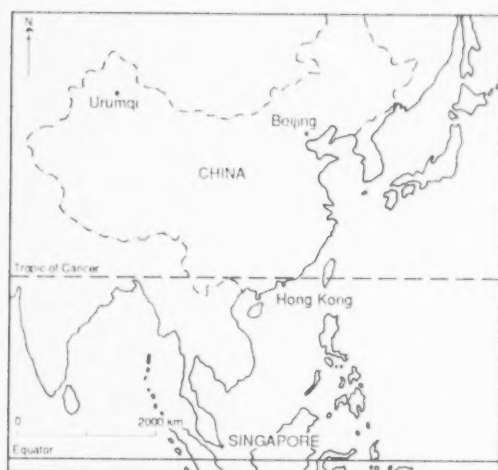


Figure 10.2 An example of a SOLO superitem in geography

The general nature of such stems in geography can be illustrated by looking at questions which are evaluated later in this chapter. The importance of graphical communication and the development of graphical skills such as map and picture interpretation is such that maps, photographs and diagrams are frequent components in ordered outcome stems in geography. In nine questions which are evaluated later in this paper, the stems consisted of:

- a simple contour map of a slope to examine the relationship between contour spacing and different slope forms;
- a simplified extract from a 1 : 20,000 Hong Kong Government map to assess understanding of simple landforms;
- a pair of large scale thematic maps showing agricultural types and the distribution of rainfall in China to examine the relationship between agriculture and climate;

- a large scale map of Asia and a set of climate graphs to examine the relationship between climatic regime, and latitude and distance from the sea;
- a map showing the location of factory areas in Hong Kong to examine the link between factory location and availability of labour and access to the port;
- a map showing the location of agricultural areas in Hong Kong to show the relationship agriculture and relief;
- a simplified diagrammatic representation of land use in Hong Kong to examine the concept of land use zones;
- a pie-gram showing the relative proportions of land use types in Hong Kong to examine land use change in Hong Kong; and
- a set of cartoons illustrating the causes of noise pollution to examine environmental factors influencing the incidence of noise pollution.

Photographs were not used purely for logistical reasons in the preparation of the test papers. Text could also have been but this might have been time consuming for the pupils and might have put too great emphasis on linguistic ability. Moreover, graphical ability is often seen to be a distinctive feature learning in geography.

10.4.3 Unistructural (U) and multistructural (M) level questions

In questions at the unistructural level pupils have to extract one piece of information from the data provided in the stem. At the multistructural level increasing complexity manifests itself in questions which require pupils to extract several linked pieces of information from the stem. The nature of these questions are now similarly illustrated by reference to selected questions from those given in the previous section.

In the question focusing on assessing understanding the relationship between slope form and contour pattern, pupils were asked to state the height of a point on a numbered contour (U) followed a question seeking pupils to differentiate between the spacing of two given contours in different parts of the map (M). In a second example drawn from the analysis of climate data, at the unistructural level (U) pupils had to read the average annual temperature from a graph on which it was highlighted; this was followed by a question comparing the number of times the average is greater at one station than at another (M). A similar style was used in the land use change question, where pupils extracted the percentage of land used for farming in 1987 (U) and then found the change in agricultural area between 1987 and 1954 (M). In a variant of this in the question on farming areas, pupils were asked to find the number of agricultural areas on Hong Kong Island which were shown (U) and to name the largest area in the territory as a whole (M). Similarly in the industrial location question, pupils had to identify from map evidence the "old" factory area (U) and then name the largest factory area which were portrayed (M). In the questions on crop pattern and climate and on urban land use in Hong Kong, unistructural questions required pupils to name the wheat growing area shown on the map and to name the land use shown in Tsim Sha Tsui respectively. At the multistructural level sequences of crops and land uses were demanded in the two cases.

The unistructural questions focus on the ability to establish single pieces of information through naming features or locations or by recognising other characteristics present in the data such as size. At the multistructural level in geography, questions emphasise aspects involving a number of facets such as relative size,

relative position, the association of a number of features; in essence they are pattern seeking. However neither of the two questions in the superitems as used are sufficient on their own in respect of the concept being investigated. Indeed there may be a problem if the same style is repeated frequently as the way of thinking may be learned as a mechanical routine and hence in reality problem solving skills may not be demonstrated. Variety is essential. They should be typical of a set of unistructural and multistructural tasks which needed for understanding of the concept. What is being attempted in the superitem is sampling of a particular level the assumption being the success would imply ability to complete other similar tasks required to develop the concept and hence understand the concept at the stated level.

10.4.4 Relational (R) level questions

These highlight interactions which provide interpretation. For example among the nine questions, pupils were asked:

- What is the relationship that can be seen between the spacing of contours and the steepness of slope shown on the map?
- What is the relationship between latitude and average annual temperature in East Asia?
- What is the relationship between relief and the location of farmland in Hong Kong?
- What changes in crops are found in China associated with change in rainfall?

Other questions may be less directly expressed as seeking relationships but they are nonetheless implicitly relational. For example, in the location of industry question, pupils were asked: "From the information on the map, what are the advantages of having a factory in Tsuen Wan rather than Yuen Long?" Similar phrasing was used in other topics: "From information on the map, what is the reason for the location of factories near Central and Tsim Sha Tsui?" or "By looking at the changes in land area under rice and under vegetables between 1954 and 1988 what changes have taken place?" Whilst there are differences in language structure and it is not certain what effect this has, all the items have in common a desire to test how far the pupils can identify linkages which are apparent within the data provided within the stem.

10.4.5 Extended abstract (E) level questions

These questions examine pupils' abilities to generalise outside the confine of the data. A number of approaches is possible here. Pupils can be asked to state underlying principles. However many principles in geography are treated in schools in an implicit rather than explicit manner. Moreover pupils may not have the language ability to express such abstract ideas directly and, if it does occur, the definition may have been rote learned. A teacher who discusses "friction of distance" with a secondary school class is likely to meet with difficulties but the class is likely to understand the costs increase with distance. It is thus important to recognise the level of abstraction which is to be expected. In many situations in geography, concepts are multi-factorial and hence one test of extended abstract level of understanding is to see whether pupils can identify factors other than those which can be derived from the stem on the basis that success indicates general conceptual

understanding. Thus in the case of the question looking at the relationship between average annual temperature and latitude pupils were asked to state and explain other factors which would influence the temperature at a place on the basis that this knowledge reflects a broader conceptualisation of climate. A similar approach was adopted in assessing the conceptualisation of the causes of land use change, the distribution of agricultural and industrial land and crop patterns at a continental scale. Another approach is to ask pupils to predict new situations from the relationship that was identified at the previous level. This was used in the problem involving the representation of slopes on contour maps; pupils were asked: "If the horizontal distance between two sets of points shown on the map were the same but the slope between the one set was twice that of the other what could you tell about the difference in height between the two locations?" In this case the pupil had to generalise the relationship between distance, height difference and gradient but, instead of working towards gradient as in the preceding relational question, must rearrange the relationship to focus on height difference. This requires the pupil to be able to conceptualise the relationship as a whole and not just mechanically apply the formula.

10.5 The validity of a SOLO format in geography

In the multiple choice question discussed in section 2, it is apparent that the subitems are at two levels of thinking. In SOLO terms they are at multistructural and relational levels. The problem in that question of not knowing whether any failure is the result of lack of learning in the topic or the difficulty of the task is serious particularly for junior secondary pupils. If we argue that using consistent SOLO levels in multiple choice questions or that formulating graded questions with a SOLO format is a way forward, then we need to be certain that the SOLO structure does in fact reflect the sort of progression of ability we see in geography classrooms. During 1989 a team from the University of Hong Kong evaluated sets of SOLO based questions from a number of school subject areas in Hong Kong. The results from this study focusing on solely the geography element are now examined. Three questions were posed, namely:

To what extent does achievement decline with increasing SOLO level within each group of pupils? (Hypothesis 1)

To what extent does achievement at a given SOLO level increase with age? (Hypothesis 2)

How do responses to superitems vary with different question styles and different materials in the stem? (Hypothesis 3)

10.5.1 Experimental design

In order to evaluate the SOLO format in the teaching of geography, nine SOLO "super-items" were tried out with two groups of secondary school pupils in Hong Kong. The pupils comprised 322 Form 1 (F1) pupils from four neighbourhood schools with a mean age of 12.7 years and 165 Form 4 (F4) pupils (mean age 15.0

years) from 3 similar schools. Each group contained a similar number of boys and girls. The nine questions, which were presented in a bilingual test format to minimise language difficulties, were constructed drawing on content areas which experience suggested were familiar to both groups being covered in both the F1 and the F4 syllabi. The questions were examined for their content validity by a panel of experienced geography educators. This was borne out by comments from the teachers in the study schools on the appropriateness of subitems for their own pupils and as to whether the content area had, or would be taught; they confirmed that, in most cases, having drawn material from commonly used Hong Kong textbooks all of which strongly conform to the centrally devised geography curriculum, the items were valid. The only superitems which raised some doubt with some teachers were 6 (on the relationship between crop patterns and climate in China), and the relational and extended abstract subitems of question 1 (on spur-valley contour patterns). To further evaluate the superitems, two pupils, aged 13 and 16 years respectively, were interviewed on completion of the tests in order to examine the problems they found in answering the questions and the way in which they came to their answers.

Fifteen class teachers were involved from the schools. They were relatively new teachers with between one and four years of experience. Most had a first degree with a teaching qualification although three of the eight F1 teachers were non-graduates. Their reported aims in teaching geography were, as might be expected varied, but both F1 and F4 teachers noted the importance of compliance with the syllabus, passing examinations, developing knowledge and skills of the subject, and developing critical thinking. A few highlighted creating awareness of daily issues and interest in the subject as important. However when asked specifically about the importance of the latter in determining what was taught on a day to day basis, most attested it was very important. Similarly most teachers saw their day to day classroom work as being directed towards the development of critical thinking and acquiring systematic knowledge. They seemed more certain of their short term than long term goals being, as Gerber (1990) found, discipline and knowledge orientated. With regard to classroom approaches, "question and answer" sessions seemed the most frequently used with all classes. "Discussion with notes" was frequent in F1 classes but "open ended approaches" were rare; on the other hand, in F4 classes the opposite was found. In using SOLO, or any level based system, interpretation is a little more difficult than when handling data from traditional right/wrong test items. In graded tests there are two dimensions, the overall score as before but also the levels achieved by pupils on different items. The partial credit model (PCM) of Masters (1982, 1988) attempts to provide a mathematical description of responses by examinees to test questions and a means of evaluating questions. It is a variant of the Rasch model and attempts to fit scores to a unidimensional latent trend. In the description, the probability of choosing a particular answer to a question is assumed to be an exponential function of two quantities: a quantity which describes the difficulty associated with identifying that particular answer, and a quantity which describes the ability of the examinee. The major assumption of the model is that the probability of choosing a particular answer to a test question depends only on the ability of the examinee and the difficulty of the answer of the test question.

Although the partial credit model can be applied to any test situation, the model is most appropriate when answers to a question can be ordered in levels of understanding. This is obviously the case when answers to a SOLO test item are marked as prestructural, unistructural, multistructural, relational or extended abstract. When

PCM is applied to data obtained from SOLO items, the model defines an item map of (i) a *difficulty* level for each of the five levels of understanding of a test question, and (ii) an *ability* for each of the examinees who took the test, expressed in the ability presumably required by higher level SOLO responding. When the ability of an examinee and the difficulty of the different levels of understanding of a particular question are known, it is possible to predict the probable level of understanding of the response by the examinee to that test question. This is given by the "logit" value. The partial credit model incorporates a notion of error in its description of the responses of examinees to the test question, which allows for the fact that some very able candidates may be unable to answer a very easy question, or that some very weak candidates may answer a very difficult question correctly. PCM thus provides a much more realistic description of actual testing situation than a deterministic model such as the Guttman scale which does not allow for careless errors made by examinees.

The model offers more than description, however; it can be used as a means of identifying the ability of an examinee or the difficulty of a level of understanding in a particular question. First, examinees are given the test questions, and their answers to each question are scored in terms of levels of understanding. Given these data, the ability of examinees and the difficulties of test answers can be estimated using standard statistical procedure such as maximum likelihood method. A computer program PC-CREDIT (Masters & Wilson 1988) is now available to carry out such an analysis. When PCM is used to determine the ability of examinees, it assigns to each examinee a value which best represents his/her performance on the test. The ability estimates for examinees are readily interpretable in that the ability ranges from negative infinity to positive infinity with large values representing high ability and small values representing low ability.

For the purposes of item screening, a t-statistic is calculated for each SOLO superitem. This acts as a measure of the fit of each SOLO based question to the PCM model and hence of the strength of the question as an ordered outcome item. The t-statistic gives an indication of how well the sequence of subquestions in a particular question performed in respect of SOLO. Significant lack of fit is generally accepted when t is either greater than 3.0 or less than -3.0. The derivation of the t-statistic is given in Appendix 1.

10.5.2 Results

Analysis was carried out in two stages. In the first, the results were examined for their reliability and the degree to which they conformed to the SOLO model (hypothesis 1). This initial work was used to screen out questions. In the second stage, the remaining questions were examined from the point of view of differences between age groups (hypothesis 2).

A. Initial screening

The "item map" from the partial credit modelling of the results provides a graphical expression of the quality of the questions as ordered outcome items. The item map is a probability plot in two dimensions. The y axis represents the scaled ability (logit) of the student whilst the x axis shows decreasing question difficulty. According to the structure of the model, pupils with low ability might be expected, at a probability of greater than 50%, either not to be able to answer the questions at all, that is in

SOLO terms answer at prestructural level, marked "0" on the item map, or to answer at a unistructural level, marked "1". Able pupils might answer more difficult multi-structural (2) or relational (3) level questions or even easy extended abstract items (4). Where items correspond well to an ordered outcome structure, levels are distributed on the item map in diagonal bands increasing in level from bottom left to top right. Table 10.1 shows the item map generated from the data in which the broad pattern is consistent with an ordered outcome structure. This can be evaluated by examining the fit of questions with acceptable reliability and by examining the % of responses which lies within the expected SOLO sequence.

Table 10.1 Results of the partial credit modelling

		DIFFICULTY									
		high ----- low									
		Question									
		(1)	8	6	(5)	3	7	(9)	2	4	
ABILITY (logit)	0.2	1	2	3	3	2	3	3	3	3	Form4
	0.1	1	2	3	3	2	3	3	3	3	
	0.0	1	1	1	3	2	3	3	3	3	
	-0.1	1	1	1	3	2	3	3	2	3	
	-0.2	1	1	1	3	2	2	3	2	3	
	-0.3	0	1	1	3	2	2	3	2	3	Form 1
	-0.4	0	1	0	3	2	2	3	2	3	
	-0.5	0	1	0	3	2	2	3	2	3	
	-0.6	0	1	0	2	1	2	3	2	3	
	-0.7	0	1	0	2	1	2	3	2	3	
-0.8	0	1	0	2	1	2	3	2	3		
-0.9	0	1	0	2	1	2	3	2	2		
Level		1 = unistructural;					2 = multistructural;				
of answer:		3 = relational;					4 = extended abstract.				

Questions which were subsequently screened out are shown in parenthesis

The degree of fit of questions to the ordered outcome structure and, hence one measure of the quality of a question, is given by a t-statistic. As noted earlier, Masters (1982) indicated that items with a t-value of greater than +3.0 or less than -3.0 are suspect. On this ground, item 9 is problematic (see table 10.2). The reliability indices (Cronbach alpha) of questions 1 and 5 are low and hence these ordered-outcome items are suspect. Furthermore a large % of responses from questions 5 and 9 were outside the expected SOLO cumulative ordered outcome pattern of no subitems correct (0000), the unistructural subitem alone (1000), the unistructural and multi-structural subitems (1100), all subitems except for extended abstract (1110), and all subitems (1111). For example some pupils answered only the unistructural and relational subitems correctly (1010) which is not consistent with model. This may be due to carelessness or the fact that relational subitems corresponded to closely to material previously taught in class, and hence the question was answered through recall. In other cases, question structure may be a problem. For example in question 5, the unistructural subitem required pupils to state the number of farming areas located on Hong Kong Island and Kowloon of which there are none. The negative answer may well have confused some children for fewer than expected F1 pupils got the answer right (41%) in comparison with the numbers correctly answering the multistructural (91%) and relational (77%) subitems. The remaining six questions (2,3,4,6,7 and 8), on which the remainder of the analysis is based, proved satisfactory on both counts with reliability values between 0.308 and 0.478, and % response patterns within the SOLO structure of between 61.3 and 83.3%. Moreover in these, the inter-correlation between sub-items in a given superitem was positive emphasising the coherence of that question. Details of these superitems are given in summary form in figure 10.3. The superitems which were screened out are given in Appendix 2.

Table 10.2 Item quality

Item	Content	t	alpha
1	Spur-valley contour patterns	1.04	0.11
2	Contour spacing and slope angle	-0.66	0.30
3	Temperature in relationship to latitude and distance from the sea	-2.89	0.43
4	Temporal change in land-use in Hong Kong	-0.59	0.31
5	Factors influencing the location of agriculture in Hong Kong	-0.34	0.14
6	Climate and crop distribution in China	-2.36	0.48
7	Factors influencing the location of factories in Hong Kong	-0.74	0.33
8	Factors influencing landuse in Hong Kong	-0.31	0.37
9	Causes of noise pollution	3.42	0.34

alpha = Cronbach alpha

B. Data analysis

Overall the results from individual supertems (table 10.3) provide evidence to support the SOLO framework for modelling pupil understanding in geography. In supertems 2, 3 and 4, a large percentage of results from both forms 1 and 4 demonstrate progressive decline in achievement with increasing level. However, supertems 6 and 8, in particular, performed less well with Form 1 pupils than with Form 4 children. In these questions, there was lower than expected numbers of correct responses at the multistructural level, or alternatively more than expected at the relational levels.

Table 10.3 Percentages of pupils answering each subitem correctly

		Form I					Form IV			Form I Form IV	
		Sch G	Sch C	Sch B	Sch A	Sch E	Sch G	Sch B	Sch F		
Q1	U	42	29	55	29	43	36	20	55	40	36
	M	15	0	11	13	3	15	11	36	9	17
	R	3	0	0	3	1	4	14	9	2	7
	E	1	0	0	0	0	1	0	5	0	1
Q2	U	89	86	92	89	87	93	94	86	89	92
	M	68	67	82	74	70	79	77	82	71	79
	R	29	50	16	53	15	44	60	73	30	52
	E	2	5	8	5	1	19	26	59	3	26
Q3	U	72	43	71	61	70	87	86	95	66	88
	M	35	19	42	47	49	67	69	77	39	70
	R	5	19	11	18	9	50	34	64	10	48
	E	5	5	0	13	1	16	3	14	4	13
Q4	U	56	81	96	71	85	94	91	95	76	94
	M	60	74	76	76	81	87	89	91	71	88
	R	55	74	63	50	67	82	69	59	61	76
	E	16	40	18	0	18	36	34	18	18	33
Q5	U	56	38	55	87	73	48	57	41	61	48
	M	59	74	37	47	43	88	89	91	53	88
	R	62	33	58	76	73	64	89	77	62	71
	E	2	0	0	0	1	2	3	9	1	3
Q6	U	22	40	26	50	38	63	66	86	32	67
	M	3	2	0	3	0	51	43	68	2	52
	R	37	40	29	50	61	64	40	73	44	60
	E	1	12	0	0	0	5	3	5	2	4
Q7	U	74	67	87	50	85	89	94	95	75	91
	M	84	93	100	97	97	93	94	100	92	94
	R	34	17	53	32	49	54	77	55	37	59
	E	0	5	3	0	1	5	17	9	1	8
Q8	U	57	45	66	66	63	75	77	68	60	74
	M	2	0	16	3	16	45	43	36	7	43
	R	16	5	18	26	24	24	51	32	18	31
	E	0	0	0	0	1	3	0	0	0	2
Q9	U	66	79	82	66	82	82	77	41	74	75
	M	66	55	79	74	76	52	74	23	70	53
	R	86	86	82	79	90	78	94	59	85	78
	E	6	29	13	18	15	13	17	23	14	15
Logit		-.71	-.51	-.45	-.37	-.37	-.49	-.36	-.28	-.53	-.33

Superitem

2. Figure 2 is a map of slopes. Study it and answer the following:
 - U What is the height of point A?
 - M The difference in height between A and B and X and Y are the same. Is the horizontal distance between A and B greater or less than that between X and Y?
 - R What is the relationship between spacing of contours and steepness of slope?
 - E If the horizontal distance between A and B and X and Y were the same, but slope between X and Y was twice that of A and B, what could you tell about the height difference?

3. Look at graphs A, B and the map (see figure 10.2).
 - U From the graph what is the average annual temperature in Singapore? (Do not calculate it)
 - M How many times greater is the average annual temperature in Singapore than in Urumqi?
 - R From the data, what is the relationship between latitude and average annual temperature?
 - E Is latitude the only factor which is likely to explain differences in temperature? Explain your answer.

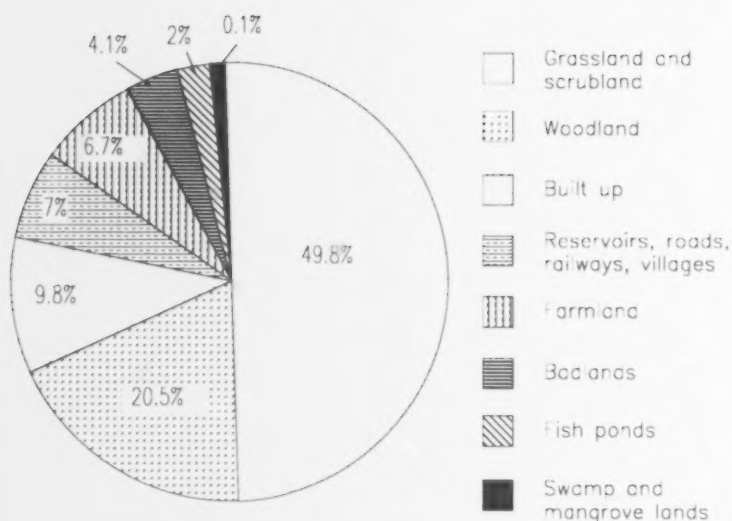
4. Look at the diagram and the table which show the land use in Hong Kong and the way it has changed (see figure 10.4).
 - U What % of land in Hong Kong is used for farming?
 - M What is the change in area with rice from 1954 to 1987?
 - R By looking at the changes in the 2 crops in table 4, what changes have taken place?
 - E Explain 2 reasons why there is less rice.

6. Look at the two maps which shows the agricultural regions and rainfall in China.
 - U Where is the main wheat growing area in China?
 - M What changes in crops is found if you travel from NE to SE China?
 - R What changes in rainfall are found with the change in crops you described in question b?
 - E Apart from rainfall what are the general conditions where wheat is grown?

7. Look at the map which shows the main industrial areas in Hong Kong.
 - U Name the old industrial area.
 - M Which is the biggest industrial area?
 - R From information on the map, what is the advantage of having a factory in Tsuen Wan rather than Yuen Long?
 - E Apart from the advantage you mentioned in question R, explain 3 other factors which influence the location of factories.

8. Look at the map which shows the pattern of landuse in Hong Kong.
 - U What is the landuse in Tsim Sha Tsui?
 - M If you travelled from Tsim Sha Tsui to Shatin along the line shown what types of landuse would you cross? Put them in order from Tsim Sha Tsui to Shatin.
 - R From information on the map what is the reason for the location of factories near Central and Tsim Sha Tsui?
 - E What are the reasons why landuse is found in different zones?

Figure 10.3 SOLO superitems used in the analysis. (Only the graphs, maps and diagrams of superitems 3 and 4 have been added)



Year	Rice	Vegetables and Flowers
1954	9450 ha	910 ha
1974	1500 ha	4560 ha
1987	10 ha	2510 ha

Figure 10.4 Land use in Hong Kong (1987)

Indeed it is likely that the problem lies at the relational level in many cases where pupils give rote learned statements of explanatory associations between geographical phenomena and environmental characteristics. If this is correct, it reflects a serious deficiency in lower secondary school geography teaching in Hong Kong. Indeed the importance of rote learning was brought out in interviews carried out with two of the pupils in their first language. They frequently commented, when they could not give relational or extended abstract answers, that they had "not learned that particular aspect". On the other hand, part of the problem may lie with the way some multistructural level subitems are framed. For example, in question 8, on spatial change in landuse in urban Hong Kong, pupils had to extract and write down a sequence of land use types and this they found difficult whereas the multistructural subitem for question 7, where they had to recognise the single biggest industrial area, they found easy. Is it perhaps that these pupils are generally used to writing only single word answers?

Further reason to believe that the SOLO ordered outcome structure matches the progressive accumulated learning abilities of the pupils is emphasised by compar-

ing the results of F1 and F4. For example, in question 3, whilst 32.3% of F1 pupils answered at only unistructural level with 25.7% and 12.8% at up to multistructural and relational levels respectively, 12.75 of F4 pupils responded at only unistructural level but 27.25 and 27.8% of this age group answered at up to the two higher levels. This pattern of fewer F4 than expected pupils remaining at lower levels but more of F4 pupils achieving higher levels is repeated across the data, illustrating the capacity of the questions to differentiate between levels of ability. Differences between F1 and F4 can also be demonstrated from the fitting the data to the Partial Credit Model. The estimated scaled probable ability levels, or logit, for Forms 1 and 4 are shown in the figure, the median value for F1 pupils being -0.53 whilst that for F4 pupils is 0.33. If we can conceive of a "mean SOLO value" within the range 0 (pre-structural) to 4 (extended abstract), then F1 pupil achievement is at a level of 1.62 (Standard deviation 1.09) whilst the corresponding form 4 level is 2.16 (Standard deviation 0.89). This suggests that a typical form 1 pupil acts over the whole range of questions at something near to a multi-structural level but on a significant number of occasions only uni-structurally. In contrast, whilst F4 pupils also exhibit a multi-structural level of understanding, many start to demonstrate relational abilities on a number of occasions.

In similar fashion it would be possible to differentiate between classes in the same age band perhaps with the prospect of tailoring the geography taught to the strengths and weaknesses of specific pupils. The data show the variability in pupils at a given school stage. Among F1 pupils, most could not correctly answer any of the items in question 6 which some teachers had thought to be an inappropriate superitem possible on grounds of the scale of analysis. In contrast, they could cope with the material and tasks in questions 3 and 8 at a uni-structural level and questions 2 and 7 at a multi-structural level. Many found question 4 the easiest with 26.4% of pupils being able to answer at up to and including a relational level. A similar variation can be found among the F4 results although they are generally one "SOLO step" higher. Thus questions 6 and 8 are answered at an uni-structural level, 2 and 3 at a multi-structural level, and 4 and 7 at a relational level. Such variations are to be expected because some tasks and ideas are more difficult than others. For example, question 8 was based on a highly abstract diagram of landuse in Hong Kong which one suspects was difficult for pupils to conceptualize despite representing their own environment. Similarly questions 3 and 6 both used small scale maps of China and involved the correlation of map data with other information either provided on a graph or on a supplementary map; these tasks were complex. In contrast, questions 2 and 4 were more straight forward requiring interpretation of a simple large scale simplified contour map in the former, and a piegram and table to show change of landuse in the latter. This simpler data which was more easily conceptualised permitted a higher level of learning.

It is interesting to compare and contrast the achievement patterns of the F1 and F4 pupils across the six questions. What is it which both can answer at a similar level and what is it that they answer differentially. Questions 2, 4 and 8 were generally handled at much the same level by both forms although pupils F4 were marginally better. However a difference between F1 and F4 of one SOLO level was apparent among questions 3, 6 and 7. These questions involved correlation of small scale maps, correlation of climate data with latitude at a continental scale and interpretation of a broadly based map showing the pattern of industry in Hong Kong respectively. Form 1 pupils could not answer any item from the first whereas form

4 pupils could respond successfully at an unistructural level. In question 6, F1 pupils were able to extract a single piece of information whereas forth formers could combine a number of elements, i.e. function multi-structurally. In the third question the form 1 - form 4 difference contrasted the multi-structural with relational answers. This gives us some insight into the changes in the geographical understanding present in lower secondary school pupils, something which requires further study.

10.6 Conclusion

A number of issues in using SOLO superitems in the assessment of learning in geography need to be addressed. It seems that a range of tasks can be set to assess learning at a given SOLO level. Indeed it is important that there is variety otherwise pupils may come to complete tasks mechanically. However, some tasks at a given level are inherently more difficult than others because for example they are more complex. However, the nature of that order of difficulty is not known with any precision. Indeed the ability of pupils to answer difficult tasks at a lower level in comparison with the ability to do easy tasks at a SOLO level higher needs examination. The effect of classroom style on pupil level of achievement also merits attention. It is argued that the SOLO structure is compatible with the What-Where-Why sequence of questions in geography. The extent to which correct answers to questions like those illustrated in this paper measure general problem solving ability or the capacity of a pupil to work geographically needs examination; the two, of course, are not independent. The way pupils arrive at their answers is an important area for further study. Some pupils may use more sophisticated strategies than others but this may not be apparent in the answer. Different routes can lead to the same observed outcome; the extent to which the use of varying paths is significant is not known and requires further investigation. Superitems provide the teacher with many advantages in respect of marking and interpretation. They also provide the pupils with a structure. However the structure may, at times, make it difficult for that level of sophistication to become readily apparent. Moreover, a complex geographical problem will usually involve a number of unistructural or multistructural type questions. Perhaps one way forward is to consider SOLO with more than two lower order subitems followed by broader more open ended higher level subitems. The results from this study support the SOLO format for testing in geography. The task specific and the data medium specific nature of the orderly cumulative progression suggested by the model, when used in a geographical context at least, has to be stressed. However this is not a major drawback if it is recognised at the outset. Testing within the model used here helps us to differentiate between topic difficulty and hence improve curriculum planning in particular in the context of mixed ability teaching. Issues remain for further research but overall headway is being made in respect of the desire to improve the nature of progression highlighted by Bennetts (1981) a decade ago.

Acknowledgements

The description of the SOLO model and the Partial Credit Model is largely based on Biggs et al. (1989). The computational assistance of Mr. Raymond Lam of the Department of Education, University of Hong Kong, is gratefully acknowledged. Responsibility for the interpretations presented here are, of course, the author's.

References

- BENNETTS, T. (1981), Progression in the geography curriculum. In: R. Walford, ed., *Signposts for geography teaching*. London: Longman.
- BIDDLE, D. (1978), A conceptual approach to teaching geography in the secondary school. *Textos Didaticos em Geografia* 2.
- BIGGS, J.B. & K. COLLIS (1982), *Evaluating the quality of learning*. New York: Academic Press.
- BIGGS, J.B. et al. (1989), Curriculum objectives and criterion referenced assessment in various secondary school subjects: A symposium. Paper at the school based innovations: looking forward to the 1990s conference. Hong Kong: University of Hong Kong.
- COLLIS, K.F. & H.A. DAVEY (1986), A technique for evaluating skills in high school science. *Journal of Research in Science Teaching* 23 pp. 651-663.
- COURTNEY, T.D. (1989), The significance of the SOLO taxonomy for learning and teaching in geography. *Geographical Education* 5, pp. 47-50.
- GERBER, R. (1990), Asian geographical educators perception and practices. In: P.G. Stimpson & Y.L. Kwan, eds., *Teaching geography in and about Asian Pacific countries. Abstracts and papers*, International Geographical Union Commission on Geographical Education Symposium. Hong Kong: University of Hong Kong.
- HAMMOND, J. (1983), *Evaluating student performance in geography*. Australia: University of Newcastle. (Unpublished B.A. Thesis).
- MARTON, F. (1981), Phenomenography -- Describing conceptions of the world around us. *Instructional Science* 10, pp. 177-200.
- MASTERS, G.N. (1982), A Rasch model for partial credit scoring. *Psychometrika* 4, pp. 149-174.
- MASTERS, G.N. (1988), Partial credit model. In: J.P. Keeves, ed., *Handbook of Educational Research Methodology, Measurement and Evaluation*. London: Pergamon Press.
- RAMSDEN, P., ed. (1988), *Improving learning: New perspectives*. London: Kogan Page.
- RHYS, W. (1972), The development of logical thinking. In: N.Graves, ed., *New movements the study and teaching of geography*. London: Temple Smith.
- WILSON, M. (1989), A comparison of deterministic and probabilistic approaches to measuring learning structures. *Australian Journal of Education* 33, pp. 127-140.
- WILSON, P. (1989), Teaching for thinking in the geography classroom. In: J. Fien, R. Gerber & P. Wilson, eds., *The geography teachers' guide to the classroom*. Brisbane: MacMillan.

Appendix 1

Derivation of the t-statistic used in the partial credit model

(Source: Masters 1982).

The fit of each item (i) to the Partial Credit Model is based on estimates of the expected value (E_{ni}), expected variance (W_{ni}) and expected kurtosis (C_{ni}) of each child's score on item i . The resultant measure follows the t distribution and is calculated by:

$$t = [(v_i^{0.33} - 1) (3q_i^{-1})] + 0.33q_i$$

where

$$v_i = [\sum (x_{ni} - E_{ni})^2] [\sum W_{ni}]^{-1}$$

$$q_i^2 = [\sum (C_{ni} - W_{ni}^2)] [\sum W_{ni}]^{-1}$$

The expected kurtosis of each child's score on item i is

$$C_{ni} = [\sum (k - E_{ni})^4] P_{kni}$$

The expected variance of each child's score on item i is

$$W_{ni} = [\sum (k - E_{ni})^2] P_{kni}$$

The expected value of a child's score on item i is

$$E_{ni} = \sum (k P_{kni})$$

where k = category (1 - 4)
 P_{kni} is the estimated probability
 of a person (n) responding correctly
 in category k to item i .

Appendix 2

SOLO based questions which were screened out:

1. Figure 1 is a contour map of a part of Hong Kong Island. Study it and answer the following questions.
 - U In the figure, do the contours in a valley "point" towards the highland, lowland or neither?
 - M Which of the sites shown by numbers on the map are spurs?
 - R If three spurs are shown next to each other on a map, what pattern of valleys would you expect?
 - E How do spurs and valleys form?

Comment:

Overall this superitem was too difficult with few correct answers beyond the U level; hence this question lacks power to differentiate between pupils. Few pupils had any notion of the concept of erosional and residual landforms needed for the E level subitem.

5. Look at the map which shows the distribution of farmland in Hong Kong.
 - U How many of the main farming areas are found on Hong Kong Island and Kowloon?
 - M Where is the largest farming area?
 - R Use the information on the map to explain the relationship between relief and the location of agriculture.
 - E Explain 2 other reasons not shown on the map which account for the location of farmland.

Comment:

The particular problem of this superitem lay with F4 pupils who failed to correctly answer the U sub-item. More than expected correctly answered the R item possibly as this was an easily recalled response which did not ask for evidence from the map.

9. Look at the diagram which shows 7 sources of noise pollution.
 - U Which of the sources 1 to 7 is most likely to be caused by the boy himself?
 - M What are the sources of noise on the diagram which come from building sites?
 - R The sources of noise on the diagram are in 3 pairs. Number 5 is on its own. Put the others in pairs.
 - E Classify the 7 pictures into 4 types of pollution.

Comment:

Overall this superitem lacked difficulty with most pupils being able to answer the U, M and R level subitems; hence this question lacks power to differentiate between pupils. The R level subitem was found to be easier than expected probably because pupils reached their decision through simple visual association rather than arguing causality. This emphasises the need to predict how pupils will arrive at answers when designing questions. Most pupils lacked the broad conceptualisation needed for the classification task in the E level subitem.

List of contributors

Prof.dr. Helmut Schrettenbrunner
University of Erlangen-Nürnberg
Regensburger Strasse 160
D-8500 Nürnberg 30
Germany
tel: 0049-911-5302523
fax: 0049-911-4010212

Prof.dr. Detlev Leutner
Abteilung Methodik
FB 06 Psychologie
Justus-Liebig-Universität
Otto-Behaghel-Strasse 10
W-6300 Giessen
Germany
tel: 0049-640-7025445
fax: 0049-640-7023811

Beate Schreiber (Dip.Psych.)
Abteilung Methodik
FB 06 Psychologie
Justus Liebig Universität
Otto Behaghel Strasse 10
W-6300 Giessen
Germany
tel: 0049-640-7025445
fax: 0049-640-7023811

Dr. Günter Hess
University of Erlangen-Nürnberg
Regensburger Strasse 160
D-8500 Nürnberg 30
Germany
tel: 0049-911-5302523
fax: 0049-911-4010212

Prof.dr. Rod Gerber
Brisbane College of Advanced Education
Queensland University of Technology
Victoria Park Road
Brisbane
Australia 4059
tel: 00617-8642111
fax: 00617-8643986

Dr. John Lidstone
Brisbane College of Advanced Education
Queensland University of Technology
Victoria Park Road
Brisbane
Australia 4059
tel: 00617-8642111
fax: 00617-8643986

Dr. Joop van der Schee
Centrum voor Educatieve Geografie
Vrije Universiteit Amsterdam
De Boelelaan 1115
1081 HV Amsterdam
The Netherlands
tel: 0031-20-5485003
fax: 0031-20-6423155

Drs. Henk van Dijk
Centrum voor Educatieve Geografie
Vrije Universiteit Amsterdam
De Boelelaan 1115
1081 HV Amsterdam
The Netherlands
tel: 0031-20-5484408
fax: 0031-20-6423155

Prof.dr. Hans van Westrhenen
Centrum voor Educatieve Geografie
Vrije Universiteit Amsterdam
De Boelelaan 1115
1081 HV Amsterdam
The Netherlands
tel: 0031-20-5487300
fax: 0031-20-6423155

Dr. Betty Lai-ling Yau
Faculty of Education
University of Hong Kong
Hong Kong
tel: 00852-8592543
fax: 00852-8585649

Dr. Kam-fai Wong
Faculty of Education
University of Hong Kong
Hong Kong
tel: 00852-8592543
fax: 00852-8585649

Dr. Lai-sim Ma
Faculty of Education
University of Hong Kong
Hong Kong
tel: 00852-8592543
fax: 00852-8585649

Prof.dr. Yee-Wang Fung
Faculty of Education
The Chinese University of Hong Kong
Shatin NT
Hong Kong
tel: 00852-6096938
fax: 00852-6036129

Prof.dr. Philip Stimpson
Faculty of Education
University of Hong Kong
Hong Kong
tel: 00852-8592543
fax: 00852-8585649

NEDERLANDSE GEOGRAFISCHE STUDIES / NETHERLANDS GEOGRAPHICAL STUDIES

- 1 G. MIK & J. H. STIKKELBROEK, Verkiezingen in Rotterdam; een geografische verkenning van de verkiezingsuitslagen 1970-1982 en een nadere analyse van de ruimtelijke structuur der Tweede Kamerverkiezingen van 1982 -- Amsterdam/Rotterdam 1985: Knag/Economisch-Geografisch Instituut Erasmus Universiteit Rotterdam. 130 pp., 51 figs., 8 tabs. ISBN 90-6809-009-7 Dfl 17,50
- 2 S. MUSTERD, Verschillende structuren en ontwikkelingen van woongebieden in Tilburg -- Amsterdam 1985: Knag/Geografisch en Planologisch Instituut van de Vrije Universiteit Amsterdam. 292 pp., 104 figs., 44 tabs. ISBN 90-6809-010-0 Dfl 27,75
- 3 M. J. TITUS, Urbanisatie, integratie en demografische respons in Jakarta; een empirisch onderzoek naar de stedelijke moderniseringsrol in de periode 1961-1976 -- Amsterdam/Utrecht 1985: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 380 pp., 14 figs., 202 tabs. ISBN 90-6809-012-7 Dfl 39,50
- 4 H. SCHENK, Views on Alleppey: Socio-historical and socio-spatial perspectives on an Industrial port town in Kerala, South India -- Amsterdam 1986: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 246 pp., 41 figs., 36 tabs. ISBN 90-6809-011-9 Dfl 29,50
- 5 P. J. BOELHOUWER & F. M. DIELEMAN (Red.), Wonen in de stad; samenvatting van sociaal-geografisch onderzoek in de Domstad -- Amsterdam/Utrecht 1986: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 138 pp., 41 figs., 32 tabs. ISBN 90-6809-013-5 Dfl 19,50
- 6 P. LUKKES & J. H. M. VAN ROODEN, De makelaardij in onroerende goederen in Nederland; een regiologisch-geografische analyse -- Amsterdam/Groningen 1986: Knag/Geografisch Instituut Rijksuniversiteit Groningen. 102 pp., 9 figs., 28 tabs. ISBN 90-6809-015-1 Dfl 25,00
- 7 P. P. P. HUIGEN, Binnen of buiten bereik? Een sociaal-geografisch onderzoek in Zuidwest-Friesland -- Amsterdam/Utrecht 1986: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 276 pp., 58 figs., 72 tabs. ISBN 90-6809-014-3 Dfl 34,00
- 8 V. M. VAN DALEN & L. VAN DER LAAN (Red.), Werken aan de kust; verslag van het Knag-symposium over de plannen tot uitbreiding van de Nederlandse kust gehouden op 3 oktober 1985 in het Provinciehuis te Den Haag -- Amsterdam 1986: Koninklijk Nederlands Aardrijkskundig Genootschap. 78 pp., 8 figs., 2 tabs. ISBN 90-6809-016-X Dfl 14,00
- 9 H. KNIPPENBERG, Deelname aan het lager onderwijs in Nederland gedurende de negentiende eeuw; een analyse van de landelijke ontwikkeling en van de regionale verschillen -- Amsterdam 1986: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 268 pp., 29 figs., 81 tabs. ISBN 90-6809-017-8 Dfl 29,00
- 10 H. J. A. BERENDSEN (Red.), Het landschap van de Bommelerwaard -- Amsterdam/Utrecht 1986: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 186 pp., 71 figs., 2 maps. ISBN 90-6809-019-4 Dfl 34,50
- 11 M. DE SMIDT (Red.), Regionale statistiek: organisatie en onderzoek; bijdragen voor drs J. Schmitz bij zijn afscheid van het CBS -- Amsterdam/Utrecht 1986: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 86 pp., 17 figs., 9 tabs. ISBN 90-6809-020-0 Dfl 14,95
- 12 J. M. VAN MOURIK, Pollen profiles of slope deposits in the Galician area (NW Spain) -- Amsterdam 1986: Knag/Fysisch-Geografisch en Bodemkundig Laboratorium Universiteit van Amsterdam. 174 pp., 55 figs., 4 tabs. ISBN 90-6809-018-6 Dfl 29,00
- 13 J. J. HARTS & L. HINGSTMAN, Verhuizingen op een rij; een analyse van individuele verhuisgeschiedenissen -- Amsterdam/Utrecht 1986: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 312 pp., 54 figs., 108 tabs. ISBN 90-6809-022-4 Dfl 38,50
- 14 A. VAN SCHAİK, Colonial control and peasant resources in Java, agricultural involution reconsidered -- Amsterdam 1986: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 214 pp., 14 figs., 31 tabs. ISBN 90-6809-021-6 Dfl 27,00
- 15 L. J. M. DIRRIX, T. K. CRIMMIUS & P. VAN DER VEEN, The functioning of periodic markets in the Bombay Metropolitan Region -- Amsterdam/Groningen 1986: Knag/Geografisch Instituut Rijksuniversiteit Groningen. 200 pp., 38 figs., 47 tabs. ISBN 90-6809-030-5 Dfl 25,00
- 16 J. C. BORCHERT, L. S. BOURNE & R. SINCLAIR (Eds.), Urban Systems in Transition -- Amsterdam/Utrecht 1986: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 248 pp., 41 figs., 48 tabs. ISBN 90-6809-028-3 Dfl 24,90
- 17 P. W. BLAUW, Suburbanisatie en sociale contacten -- Amsterdam/Rotterdam 1986: Knag/Faculteit der Economische Wetenschappen Erasmus Universiteit Rotterdam. 168 pp., 68 tabs. ISBN 90-6809-024-0 Dfl 25,00
- 18 H. J. SCHOLTEN, R. J. VAN DE VELDE & P. PADDING, Doorstroming op de Nederlandse woningmarkt; geanalyseerd en gemodelleerd -- Amsterdam/Utrecht 1986: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 116 pp., 38 figs., 22 tabs. ISBN 90-6809-025-9 Dfl 13,00
- 19 F. M. DIELEMAN, A. W. P. JANSEN & M. DE SMIDT (Red.), Metamorfose van de stad; recente tendenzen van wonen en werken in Nederlandse steden -- Amsterdam/Utrecht 1986: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 134 pp., 31 figs., 22 tabs. ISBN 90-6809-026-7 Out of print
- 20 E. VOS, M. NIEUWENHUIS, M. HOOGENDOORN & A. SENDERS, Vele handen ...; vrouw en werk in Latijns Amerika -- Amsterdam 1986: Knag/Geografisch en Planologisch Instituut van de Vrije Universiteit. 210 pp., 9 figs., 7 tabs. ISBN 90-6809-027-5 Dfl 30,00
- 21 J. H. J. VAN DINTEREN & H. W. TER HART (Red.), Geografie en kantoren 1985; verslag van een congres gehouden te Nijmegen op 29 november 1985 -- Amsterdam/Nijmegen 1986: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 144 pp., 15 figs., 15 tabs. ISBN 90-6809-029-1 Dfl 17,00
- 22 J. VIJGEN & R. v. ENCELSDORP GASTELAARS, Stedelijke bevolkingskategorien in opkomst, stijlen en strategieën in het alledaags bestaan -- Amsterdam 1986: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 122 pp., 3 figs., 40 tabs. ISBN 90-6809-031-3 Dfl 15,00

- 23 H J MÜCHER, Aspects of loess and loess-derived slope deposits; an experimental and micromorphological approach -- Amsterdam 1986: Knag/Fysisch-Geografisch en Bodenkundig Laboratorium Universiteit van Amsterdam. 268 pp, 42 figs, 9 tabs. ISBN 90-6809-032-1 Dfl 25,00
- 24 P HENDRIKS, De relationele definitie van begrippen; een relationeel realistische visie op het operationaliseren en representeren van begrippen -- Amsterdam/Nijmegen 1986: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 282 pp, 28 figs, 7 tabs. ISBN 90-6809-033-X Dfl 30,00
- 25 J M G KLEINPENNING (ed), Competition for rural and urban space in Latin America; its consequences for low income groups. Contributions to a symposium organized at the 45th International Congress of Americanists, Bogota, 1-7 July 1985 -- Amsterdam/Nijmegen 1986: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 178 pp, 36 figs, 11 tabs. ISBN 90-6809-034-8 Dfl 22,50
- 26 J BUURSINK & E WEVER (red), Regio en ontwikkeling; aspecten van regionaal-economische ontwikkelingen in Nederland -- Amsterdam/Nijmegen 1986: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 160 pp, 41 figs, 50 tabs. ISBN 90-6809-035-6 Dfl 20,00
- 27 G CLARK, P DOSTAL & F THISSEN (eds), Rural research and planning: the Netherlands and Great Britain. Report of the Second British-Dutch Symposium on Rural Geography and Planning, 5 September 1986 -- Amsterdam 1987: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 88 pp, 6 figs, 4 tabs. ISBN 90-6809-037-2 Dfl 10,00
- 28 W M KARREMAN & M de SMIDT (red), Redevoeringen en kleine geschriften van Prof A C de Vooy; verzameling van niet eerder uitgegeven werk ter gelegenheid van zijn tachtigste verjaardag, 22 jan. 1987 -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 156 pp, 8 figs, 5 tabs. ISBN 90-6809-036-4 Dfl 21,70
- 29 G PEPPERKAMP (red), Mens en milieu in de derde wereld; bijdragen aan de Nederlandse Geografendagen te Utrecht, 2-4 april 1986 -- Amsterdam/Nijmegen 1987: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 146 pp, 17 figs, 11 tabs. ISBN 90-6809-038-0 Dfl 20,00
- 30 A R WOLTERS & A PERSMA, Beschermde reservaten? Een milieugeografische benadering -- Amsterdam/Groningen 1987: Knag/Geografisch Instituut Rijksuniversiteit Groningen. 184 pp, 47 figs, 4 tabs. ISBN 90-6809-039-9 Out of print
- 31 W J v d BREMEN & P H PELLENBARG (red), Het geografisch plechtanker: eenheid in verscheidenheid. Liber amicorum Rob Tamsa -- Amsterdam/Groningen 1987: Knag/Geografisch Instituut Rijksuniversiteit Groningen. 336 pp, 58 figs, 22 tabs. ISBN 90-6809-040-2 Dfl 35,00
- 32 G MIK, Segregatie in het grootstedelijk milieu; theorie en Rotterdamse werkelijkheid -- Amsterdam/Rotterdam 1987: Knag/Economisch-Geografisch Instituut Erasmus Universiteit Rotterdam. 252 pp, 48 figs, 45 tabs. ISBN 90-6809-041-0 Dfl 25,00
- 33 H J M GOVERDE, Macht over de Markerruimte -- Amsterdam/Nijmegen 1987: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 480 pp, 26 figs, 22 tabs. ISBN 90-6809-042-9 Dfl 57,50
- 34 P P GROENEWEGEN, J P MACKENBACH & M H STIJNENBOSCH (red), Geografie van gezondheid en gezondheidszorg -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 112 pp, 25 figs, 19 tabs. ISBN 90-6809-043-7 Dfl 19,70
- 35 R ter BRUGGE & E WEVER (red), Energiebeleid, het Nederlandse energiebeleid in ruimtelijk perspectief. Verslag van een Knag-symposium -- Amsterdam/Groningen/Nijmegen 1987: Knag/Geografisch Instituut Rijksuniversiteit Groningen/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 112 pp, 21 figs, 18 tabs. ISBN 90-6809-044-5 Dfl 18,00
- 36 J A v d SCHEE, Kijk op kaarten; een empirisch onderzoek naar het gebruik van geografische denkvaardigheden bij het analyseren van kaarten door leerlingen uit het vierde leerjaar van het VWO -- Amsterdam 1987: Knag/Geografisch en Planologisch Instituut van de Vrije Universiteit Amsterdam. 112 pp, 42 figs, 58 tabs. ISBN 90-6809-045-3 Dfl 39,50
- 37 O VERKOREN & J v WEESEP (eds), Spatial mobility and urban change -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 180 pp, 17 figs, 45 tabs. ISBN 90-6809-051-8 Dfl 24,75
- 38 M W de JONG, New economic activities and regional dynamics -- Amsterdam 1987: Knag/Economisch-Geografisch Instituut Universiteit van Amsterdam. 200 pp, 26 figs, 27 tabs. ISBN 90-6809-046-1 Dfl 29,00
- 39 A C M JANSEN, Bier in Nederland en België; een geografie van de smaak -- Amsterdam 1987: Knag/Economisch-Geografisch Instituut Universiteit van Amsterdam. 282 pp, 14 figs, 7 tabs. ISBN 90-6809-047-X Dfl 37,50
- 40 Y C J BROUWERS, M C DEURLOO & L de KLERR, Selectieve verhuisbewegingen en segregatie; de invloed van de etnische samenstelling van de woonomgeving op verhuisgedrag -- Amsterdam 1987: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 112 pp, 9 figs, 22 tabs. ISBN 90-6809-048-8 Dfl 16,00
- 41 R J SCHOUW & F M DIELEMAN, Echtscheiding en woningmarkt; een voorstudie naar de complexe relatie tussen echtscheiding en woningmarkt -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 98 pp, 8 figs, 21 tabs. ISBN 90-6809-049-6 Dfl 14,95
- 42 J G GROENENDIJK, De positie van dorpen in het beleid van Nederlandse plattelandsgemeenten; een politiek-geografisch onderzoek naar de verdeling van woningbouwlokaties tijdens een fase van omvangrijke groei -- Amsterdam 1987: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 314 pp, 22 figs, 55 tabs. ISBN 90-6809-050-X Dfl 31,50
- 43 J G BORCHERT & J BUURSINK (red), Citymarketing en geografie -- Amsterdam/Nijmegen 1987: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 172 pp, 32 figs, 14 tabs. ISBN 90-6809-052-6 Dfl 20,00
- 44 J J M ANGENT & A BONGENAAR (eds), Planning without a passport: the future of European spatial planning -- Amsterdam 1987: Knag/Stichting Interuniversitair Instituut voor Sociaal-Wetenschappelijk Onderzoek. 184 pp, 26 figs, 7 tabs. ISBN 90-6809-053-4 Out of print

- 45 R C v d MARK, A H PERRELS & J J REYNDERS, Kansen voor het Noorden; een beleidsstrategisch onderzoek naar nieuwe technologie -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht/Economische Faculteit Vrije Universiteit Amsterdam. 168 pp, 54 figs, 41 tabs. ISBN 90-6809-054-2 Dfl 22,50
- 46 J J STERKENBURG, Rural development and rural development policies: cases from Africa and Asia -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 196 pp, 13 figs, 14 tabs. ISBN 90-6809-055-0 Dfl 26,50
- 47 C CORTIE, Alkmaar, van streekcentrum naar groeiern; een onderzoek naar migratie en forensisme tijdens de transformatie van een stedelijk systeem -- Amsterdam 1987: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 204 pp, 28 figs, 39 tabs. ISBN 90-6809-056-9 Dfl 25,00
- 48 J A A M KOK & P H PELLENBARG (red), Buitenlandse bedrijven in Nederland; werving in Europees perspectief -- Amsterdam/Groningen 1987: Knag/Geografisch Instituut Rijksuniversiteit Groningen. 112 pp, 17 figs, 30 tabs. ISBN 90-6809-059-3 Out of print
- 49 T DIETZ, Pastoralists in Dire Straits; survival strategies and external interventions in a semi-arid region at the Kenya/Uganda border: Western Pokot, 1900-1986 -- Amsterdam 1987: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 332 pp, 34 figs, 66 tabs. ISBN 90-6809-057-7 Dfl 43,00
- 50 F J J H v HOORN, Onder anderen; effecten van de vestiging van Mediterraan in naorlogse wijken Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 226 pp, 36 figs, 55 tabs. ISBN 90-6809-060-7 Dfl 29,70
- 51 M J DIJST & C CORTIE, Universiteit en revitalisering -- Amsterdam 1987: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 140 pp, 6 figs, 13 tabs. ISBN 90-6809-058-5 Dfl 17,00
- 52 Planologie als kleurbeepaling; de rol van toonaangevende instellingen en bedrijven op de ontwikkeling van de Amsterdamse Museum- en Concertgebouwuurt -- Amsterdam 1987: Knag/Centrum Beleidsadviserend Onderzoek. 164 pp, 2 figs, 23 tabs. ISBN 90-6809-061-5 Dfl 25,00
- 53 J VERHORST & M H STIJNBOSCH, Bedrijvigheid en stadsvernieuwing; analyse van de bedrijvigheidsontwikkeling in enkele stadsvernieuwinggebieden in Utrecht en Den Haag in de periode 1973/1974 - 1983/1984 -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 112 pp, 47 figs, 25 tabs. ISBN 90-6809-063-1 Dfl 15,70
- 54 B G J DRIESSEN, R VERHOEF & J C P ter WELLE-HEETHUIS, Overheid en bevolkingsontwikkelingen; een onderzoek naar autonome en niet-autonome bevolkingsontwikkelingen in de stadsgewesten Arnhem en Utrecht -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 166 pp, 53 figs, 42 tabs. ISBN 90-6809-064-X Dfl 23,30
- 55 O A L C ATZEMA, P P P HUIGEN, A G A de VOCHT & C R VOLKERS, De bereikbaarheid van voorzieningen in Noord Nederland -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 220 pp, 49 figs, 122 tabs. ISBN 90-6809-065-8 Dfl 24,00
- 56 P C DEURENKAMP, G A HOEKVELD & A MUDDE (red), Geografie en onderwijs televisie -- Amsterdam/Utrecht 1987: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 222 pp, 29 figs, 6 tabs. ISBN 90-6809-066-6 Dfl 26,50
- 57 G CARDOL, Ruimte voor agribusiness-complexen; structuur, positie en dynamiek van het Noordlimburgse tuinbouwcomplex vanuit functioneel, geografisch en regionaal perspectief -- Amsterdam/Nijmegen 1988: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 312 pp, 34 figs, 57 tabs. ISBN 90-6809-067-4 Dfl 30,00
- 58 M JANSEN-VERBEKE, Leisure, recreation and tourism in inner cities; an explorative study -- Amsterdam/Nijmegen 1988: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 316 pp, 61 figs, 51 tabs. ISBN 90-6809-068-2 Out of print
- 59 A H H M KEMPERS-WARMERDAM, Vergrijzen in het groen; het bereik van ouderen en de bereikbaarheid van voorzieningen in landelijke gebieden -- Amsterdam/Utrecht 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 236 pp, 47 figs, 70 tabs. ISBN 90-6809-069-0 Dfl 29,50
- 60 P J BOELHOUWER, De verkoop van woningwetwoningen; de overdracht van woningwetwoningen aan bewoners en de gevolgen voor de volkshuisvesting -- Amsterdam/Utrecht 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 208 pp, 49 figs, 122 tabs. ISBN 90-6809-070-4 Dfl 29,30
- 61 A G J DIETVORST & M C JANSEN-VERBEKE, De binnenstad: kader van een sociaal perpetuum mobile; een literatuurstudie naar tijdsbesteding en binnenstadsgebruik -- Amsterdam/Nijmegen 1988: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 240 pp, 10 tabs. ISBN 90-6809-071-2 Dfl 30,00
- 62 H SCHRETTENBRUNNER & J v WESTRHENEN, Empirische Forschung und Computer im Geographie-unterricht; Niederländisch-deutsches Symposium Amsterdam 1987 -- Amsterdam 1988: Knag/Centrum voor Educatieve Geografie Vrije Universiteit. 120 pp, 27 figs. ISBN 90-6809-072-0 Dfl 15,00
- 63 H J A BERENDSEN & H v STEIJN (red), Nieuwe karteringsmethoden in de fysische geografie -- Amsterdam/Utrecht 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 176 pp, 56 figs, 24 tabs. ISBN 90-6809-073-9 Dfl 22,50
- 64 A G J DIETVORST & J P M KWAAD (eds), Geographical research in the Netherlands 1978-1987 -- Amsterdam 1988: Knag/International Geographical Union Netherlands. 262 pp, 7 figs, 2 tabs. ISBN 90-6809-074-7 Dfl 33,00
- 65 J v WEESEP, Appartementenrechten; het gebruik van het splitsingsregime -- Amsterdam/Utrecht 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 94 pp, 4 figs, 16 tabs. ISBN 90-6809-075-5 Dfl 14,50
- 66 T W A EPPINK, Choice of mathematical models in geographic research considering alternatives -- Amsterdam/Nijmegen 1988: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 244 pp, 74 figs, 49 tabs. ISBN 90-6809-076-3 Dfl 30,00
- 67 J HINDERINK & E SZULC-DABROWIECKA (eds), Successful rural development in Third World Countries -- Amsterdam/Utrecht 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 256 pp, 14 figs, 20 tabs. ISBN 90-6809-077-1 Dfl 31,50

- 68 S BAREND, J D H HARTEN, J RENES, J VERHORST & K E v d WIELEN (red), *Planning in het verleden* -- Amsterdam/Utrecht 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 192 pp, 71 figs. ISBN 90-6809-078-X Dfl 26,00
- 69 J MANSVELT BECK, *The rise of a subsidized periphery in Spain* -- Amsterdam 1988: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 286 pp, 15 figs, 28 tabs. ISBN 90-6809-079-8 Dfl 37,50
- 70 S SMITH, *Kleinschalige industrie in Latijns Amerika; een studie van de ontwikkelingsmogelijkheden van de 'informele' kleding- en textielnijverheid in Aguascalientes, Mexico* -- Amsterdam/Nijmegen 1988: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 422 pp, 4 figs, 16 tabs. ISBN 90-6809-080-1 Dfl 42,50
- 71 W DWARKASING, D HANEMAAYER, M de SMIDT & P P TORDOIR, *Ruimte voor hoogwaardige kantoren; onderzoek naar toplocaties voor de commerciële kantorensector gezien vanuit de optiek van de gebruikers* -- Amsterdam/Utrecht/Leiden/Delft 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht/Research voor Beleid/Inro-tno. 112 pp, 7 figs, 44 tabs. ISBN 90-6809-081-X Out of print
- 72 P J KORTEWEG, *Dynamiek en immobiliteit in naoorlogse wijken; het functioneren van woonwijken in Alkmaar, Haarlem en Purmerend* -- Amsterdam/Utrecht 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 144 pp, 21 figs, 34 tabs. ISBN 90-6809-082-8 Dfl 20,90
- 73 P J WIJERS, *Land prices in Tokyo; causes and effects, government policies and implications for the real estate industry* -- Amsterdam 1988: Knag/Economisch-Geografisch Instituut Universiteit van Amsterdam. 84 pp, 12 figs, 8 tabs. ISBN 90-6809-084-4 Dfl 47,50
- 74 J v MOURIK (red), *Landchap in beweging; ontwikkeling en bewoning van een stuifzandgebied in de Kempen* -- Amsterdam 1988: Knag/Faculteit Ruimtelijke Wetenschappen Universiteit van Amsterdam. 197 pp, 95 figs, 1 tab. ISBN 90-6809-083-6 Dfl 30,00
- 75 W J v OSTENDORF, *Het sociaal profiel van de gemeente; woonmilieudifferentiatie en de vorming van het stadsgewest Amsterdam: het ruimtelijk beleid van een achttal gemeenten na de tweede wereldoorlog* -- Amsterdam 1988: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 192 pp, 12 figs, 26 tabs. ISBN 90-6809-085-2 Dfl 23,00
- 76 J de BRUIN & J A KOETSIER (red), *De kracht van de regio; sociaal-economische ontwikkelingsmogelijkheden van de regio* -- Amsterdam 1988: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 104 pp, 12 figs, 6 tabs. ISBN 90-6809-086-0 Dfl 15,00
- 77 A G M v d SYACT & P H J HENDRIKS (red), *Methoden op een keerpunt; opstellen aangeboden aan prof drs P J W Kouwe bij zijn afscheid als hoogleraar aan de Katholieke Universiteit Nijmegen* -- Amsterdam/Nijmegen 1988: Knag/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 170 pp, 20 figs, 30 tabs. ISBN 90-6809-087-9 Dfl 25,00
- 78 C v d POST, *Vagrants and migrant labour absorption in large and small centres in Swaziland; a comparative study of the towns Manzini and Nhlanguano* -- Amsterdam/Utrecht 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 310 pp, 32 figs, 84 tabs. ISBN 90-6809-088-7 Dfl 35,00
- 79 L J de HAAN, *Overheid en regionale integratie van de savanne in Togo 1885-1985* -- Amsterdam 1988: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 304 pp, 31 figs, 65 tabs. ISBN 90-6809-089-5 Dfl 33,00
- 80 L H v WIJNCAARDEN-BAKKER & J J M v d MEER (eds), *Spatial sciences, research in progress: Proceedings of the symposium "Spatial sciences, research in progress", 14 April 1988 Amsterdam* -- Amsterdam 1988: Knag/Faculteit Ruimtelijke Wetenschappen Universiteit van Amsterdam. 112 pp, 16 figs, 2 tabs. ISBN 90-6809-091-7 Dfl 24,00
- 81 F M H M DRIESSEN & J H v HOUWELINGEN, *Vrije tijd en korte verblijfsrecreatie* -- Amsterdam/Utrecht 1988: Knag/Bureau Driessen. 256 pp, 25 figs, 146 tabs. ISBN 90-6809-095-X Dfl 15,00
- 82 G HOEKVELD-MEIJER & G J SCHUTTE, *Aardrijkskunde gebiedenderwijs; tekst en uitleg bij het schrijven, lezen, denken en leren over gebieden en verschijnselen in gebieden* -- Amsterdam 1988: Knag/Centrum voor Educatieve Geografie Vrije Universiteit Amsterdam. 252 pp, 76 figs, 12 tabs. ISBN 90-6809-090-9 Dfl 39,00
- 83 P K DOORN, *Social structure and spatial mobility: composition and dynamics of the Dutch labour force* -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 262 pp, 72 figs, 41 tabs. ISBN 90-6809-092-5 Dfl 31,50
- 84 A LOEVE, *Buitenlandse ondernemingen in regionaal perspectief, vestigingsstrategieën en regionale effecten van buitenlandse bedrijven in Nederland* -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 272 pp, 49 figs, 78 tabs. ISBN 90-6809-093-3 Dfl 32,00
- 85 D H de BAKKER, *Ruraal nederzettingenpatroon en beleid; ontwikkelingen in Zuidwest-Friesland* -- Amsterdam/Utrecht 1988: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 230 pp, 32 figs, 68 tabs. ISBN 90-6809-094-1 Dfl 29,00
- 86 L J PAUL (ed), *Post-war development of regional geography; with special attention to the United Kingdom, Belgium, and the Netherlands* -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 88 pp, 15 figs, 5 tabs. ISBN 90-6809-096-8 Dfl 14,00
- 87 P HOEKSTRA, *River outflow, depositional processes and coastal morphodynamics in a monsoon-dominated deltaic environment, East Java, Indonesia* -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 220 pp, 77 figs, 24 tabs. ISBN 90-6809-097-6 Dfl 28,50
- 88 E LENSINK, *Intermediaire diensten in landelijke gebieden; een economisch-geografisch onderzoek in een rurale omgeving* -- Amsterdam/Nijmegen 1989: Knag/Faculteit Beleidswetenschappen Katholieke Universiteit Nijmegen. 246 pp, 21 figs, 65 tabs. ISBN 90-6809-098-4 Dfl 30,00
- 89 P P P HUIGEN & M C H M v d VELDEN (red), *De achterkant van verstedelijkt Nederland; de positie en functie van landelijke gebieden in de Nederlandse samenleving* -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 181 pp, 25 figs, 46 tabs. ISBN 90-6809-100-X Dfl 25,00

- 90 J H J VAN DINTEREN, Zakelijke diensten en middelgrote steden, een vergelijkend onderzoek naar de vestigingsplaatskeuze en het functioneren van zakelijke dienstverleningsbedrijven in Noord-Brabant, Gelderland en Overijssel -- Amsterdam/Nijmegen 1989: Knag/Faculteit der Beeldwetenschappen Katholieke Universiteit Nijmegen. 312 pp, 28 figs, 84 tabs. ISBN 90-6809-099-2 Dfl 40,00
- 91 L VAN DER LAAN, H SCHOLTEN & G A VAN DER KNAAP, Het regionaal arbeidsaanbod in Nederland; structuur en ontwikkeling -- Amsterdam/Rotterdam 1989: Knag/Economisch-Geografisch Instituut Erasmus Universiteit Rotterdam. 128 pp, 27 figs, 28 tabs. ISBN 90-6809-101-8 Dfl 17,50
- 92 C CLARK, P HUIGEN & F THISSEN (eds), Planning and the future of the countryside: Great Britain and the Netherlands; proceedings of the second British-Dutch Symposium on Rural Geography, 6-8 september 1986 Amsterdam -- Amsterdam 1989: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 240 pp, 25 figs, 43 tabs. ISBN 90-6809-102-6 Dfl 35,00
- 93 J A VAN DEN BERG, Variability of parameters for modelling soil moisture conditions; studies on loamy to silty soils on marly bedrock in the Ardèche drainage basin, France -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 214 pp, 76 figs, 16 tabs. ISBN 90-6809-103-4 Dfl 28,50
- 94 O VERKOREN, Huizen op de hoogvlakte; een residentieel-geografische verkenning van La Paz, Bolivia -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 210 pp, 29 figs, 16 tabs. ISBN 90-6809-104-2 Dfl 32,00
- 95 G MIK (red), Herstructurering in Rotterdam; modernisering en internationalisering en de Kop van Zuid -- Amsterdam/Rotterdam 1989: Knag/Economisch Geografisch Instituut Erasmus Universiteit Rotterdam. 324 pp, 86 figs, 54 tabs. ISBN 90-6809-105-0 Dfl 30,00
- 96 P BEEKMAN, P VAN LINDERT, J POST & W PRINS, Huisvestingsbeleid en informele bouw in de derde wereld; tussen idee en realiteit -- Amsterdam 1989: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 174 pp, 9 figs, 25 tabs. ISBN 90-6809-106-9 Dfl 30,00
- 97 J G L PALTE, Upland farming on Java, Indonesia; a socio-economic study of upland agriculture and subsistence under population pressure -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 256 pp, 15 figs, 38 tabs. ISBN 90-6809-107-7 Dfl 30,50
- 98 P VAN GENUCHTEN, Movement mechanisms and slide velocity variations of landslides in varved clays in the French Alps -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 160 pp, 70 figs, 17 tabs. ISBN 90-6809-108-5 Dfl 25,00
- 99 M DE SMIDT & E WEVER (eds), Regional and local economic policies and technology -- Amsterdam/Utrecht/Nijmegen 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht/Geografisch en Planologisch Instituut Katholieke Universiteit Nijmegen. 156 pp, 53 figs, 36 tabs. ISBN 90-6809-109-3 Dfl 24,00
- 100 P J H RIEMENS, On the foreign operations of third world firms -- Amsterdam 1989: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 148 pp, 20 tabs. ISBN 90-6809-110-7 Dfl 30,00
- 101 C B M PEDROLI, The nature of landscape; a contribution to landscape ecology and ecohydrology with examples from the Strijper Aa landscape, Eastern Brabant, the Netherlands -- Amsterdam 1989: Knag/Fysisch-Geografisch en Bodemkundig Laboratorium Universiteit van Amsterdam. 164 pp, 43 figs, 18 tabs. ISBN 90-6809-111-5 Dfl 25,00
- 102 H LEENAERS, The dispersal of metal mining wastes in the catchment of the river Geul, Belgium-the Netherlands -- Amsterdam/Utrecht 1989: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 230 pp, 95 figs, 52 tabs. ISBN 90-6809-112-3 Dfl 30,00
- 103 C A HOEKVELD, G SCHOENMAKER & J VAN WESTRHENEN, Wijkende grenzen -- Amsterdam 1989: Knag/Centrum voor Educatieve Geografie Vrije Universiteit Amsterdam. 216 pp, 38 figs, 24 tabs. ISBN 90-6809-113-1 Dfl 27,50
- 104 P C J DRUIJVEN, Mandenvlechters en Mexcalstokers in Mexico -- Amsterdam 1990: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 294 pp, 21 figs, 55 tabs. ISBN 90-6809-114-X Dfl 38,50
- 105 W BLEUTEN, De verwatering van meststoffen; analyse en modellering van de relaties tussen landgebruik en waterkwaliteit in het stroomgebied van de Langbroeker Wetering -- Amsterdam/Utrecht 1990: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 262 pp, 80 figs, 31 tabs. ISBN 90-6809-115-8 Dfl 37,00
- 106 J VAN WEESEP & P KORCELLI (eds), Residential mobility and social change; studies from Poland and the Netherlands -- Amsterdam/Utrecht 1990: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 182 pp, 46 figs, 34 tabs. ISBN 90-6809-116-6 Dfl 29,50
- 107 M VAN HERWIJNEN, R JANSSEN & P RIETVELD, Herbestemming van landbouwgrond; een multicriteria benadering -- Amsterdam 1990: Knag/Instituut voor Milieuvraagstukken Vrije Universiteit. 110 pp, 33 figs, 12 tabs. ISBN 90-6809-117-4 Dfl 25,00
- 108 D H DRENTH, De informatica-sector in Nederland tussen rijp en groen; een ruimtelijk-economische analyse. Amsterdam/Nijmegen 1990: Knag/Faculteit der Beeldwetenschappen Katholieke Universiteit Nijmegen. 268 pp, 24 figs, 87 tabs. ISBN 90-6809-118-2 Dfl 37,50
- 109 H KNOL & W MANSHANDEN, Functionele samenhang in de noordvleugel van de Randstad -- Amsterdam/Utrecht 1990: Knag/Economisch-Geografisch Instituut Universiteit van Amsterdam/Geografisch Instituut Rijksuniversiteit Utrecht. 112 pp, 26 figs, 27 tabs. ISBN 90-6809-119-0 Dfl 19,50
- 110 C D EYSBERG, The Californian wine economy: natural opportunities and socio-cultural constraints - a regional geographic analysis of its origins and perspectives -- Amsterdam/Utrecht 1990: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 272 pp, 64 figs, 26 tabs. ISBN 90-6809-121-2 Dfl 29,75
- 111 J W A DIJKMANS, Aspects of geomorphology and thermoluminescence dating of cold climate eolian sands -- Amsterdam/Utrecht 1990: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 256 pp, 119 figs, 19 tabs. ISBN 90-6809-120-4 Dfl 36,50
- 112 H TER HEIDE (ed), Technological change and spatial policy -- Amsterdam/Utrecht 1990: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 218 pp, 9 figs, 31 tabs. ISBN 90-6809-122-0 Dfl 29,00

- 113 I L M VAN HEES, De ontwikkeling van een woningmarktmodel en zijn toepassing op Italië; The building of a housing-market model and its application to Italy -- Amsterdam/Nijmegen 1990: Knag/Faculteit der Beelidswetenschappen Katholieke Universiteit Nijmegen. 196 pp, 12 figs, 26 tabs. ISBN 90-6809-123-9 Dfl 35,00
- 114 M R HENDRIKS, Regionalisation of hydrological data: Effects of lithology and land use on storm runoff in east Luxembourg -- Amsterdam/Utrecht 1990: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 174 pp, 23 figs, 50 tabs. ISBN 90-6809-124-7 Dfl 26,00
- 115 P H RENOODY, The informal economy: meaning, measurement and social significance -- Amsterdam 1990: Knag/Regioplan. 204 pp, 12 figs, 21 tabs. ISBN 90-6809-125-5 Dfl 35,00
- 116 J H T KRAMER, Luchthavens en hun uitstraling; een onderzoek naar de economische en ruimtelijke uitstralingseffecten van luchthavens -- Amsterdam/Nijmegen 1990: Knag/Faculteit der Beelidswetenschappen Katholieke Universiteit Nijmegen. 312 pp, 47 figs, 60 tabs. ISBN 90-6809-126-3 Dfl 55,00
- 117 M DE KWAASTENIET, Denomination and primary education in the Netherlands (1870-1984): a spatial diffusion perspective -- Amsterdam/Florence 1990: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam/European University Institute Florence. 268 pp, 28 figs, 39 tabs. ISBN 90-6809-127-1 Dfl 36,00
- 118 W P M F IVENS, Atmospheric deposition onto forests: An analysis of the deposition variability by means of throughfall measurements -- Amsterdam/Utrecht 1990: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 156 pp, 53 figs, 36 tabs. ISBN 90-6809-128-X Dfl 25,00
- 119 R HASSINK, Herstructurering en innovatiebevordering in het Ruhrgebied; een onderzoek naar de beleids- en bedrijfsreacties op de herstructureringsproblematiek -- Amsterdam/Utrecht 1990: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 122 pp, 20 figs, 18 tabs. ISBN 90-6809-120-8 Dfl 24,00
- 120 P P SCHOT, Solute transport by groundwater flow to wetland ecosystems; The environmental impact of human activities -- Amsterdam/Utrecht 1991: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 136 pp, 27 figs, 9 tabs. ISBN 90-6809-130-1 Dfl 25,00
- 121 S DEN HENGST & B DE PATER (Red), Externe relaties en regionale ontwikkeling: voorbeelden uit Spanje en Portugal -- Amsterdam/Utrecht 1991: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 198 pp, 31 figs, 27 tabs. ISBN 90-6809-131-X Dfl 29,50
- 122 J KROES, Onvolledige opstreng op de Nederlandse zandgronden; een onderzoek naar de verspreiding en de achtergronden van overgangsvormen tussen opstreng en andere occupatievormen -- Amsterdam/Utrecht 1991: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 256 pp, 65 figs. ISBN 90-6809-132-8 Dfl 35,00
- 123 H S VERDUIN-MULLER, Serving the knowledge-based society: research on knowledge products -- Amsterdam/Utrecht 1991: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 116 pp, 3 figs. ISBN 90-6809-133-6 Dfl 24,00
- 124 F MULDER, Assessment of landslide hazard -- Amsterdam/Utrecht 1991: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 156 pp, 59 figs, 25 tabs. ISBN 90-6809-134-4 Dfl 29,50
- 125 M VIS, Processes and patterns of erosion in natural and disturbed Andean forest ecosystems -- Amsterdam 1991: Knag/Fysisch Geografisch en Bodemkundig Laboratorium Universiteit van Amsterdam. ca 158 pp. ISBN 90-6809-136-0 Dfl 12,00
- 126 V EIFF, Beleid voor bedrijfsterreinen; een politiek-geografische studie naar betrekkingen in het openbaar bestuur bij het lokatiebeleid voor de bedrijfsvestiging in stadsgewesten -- Amsterdam 1991: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 214 pp, 13 figs, 18 tabs. ISBN 90-6809-135-2 Dfl 37,50
- 127 O ATZEMA, Stad uit, stad in; residentiële suburbanisatie in Nederland in de jaren zeventig en tachtig -- Amsterdam/Utrecht 1991: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 274 pp, 57 figs, 80 tabs. ISBN 90-6809-137-9 Dfl 37,50
- 128 M HULSHOF, Zatopec moves; networks and remittances of US-bound migrants from Oaxaca, Mexico -- Amsterdam 1991: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 106 pp, 6 figs, 14 tabs. ISBN 90-6809-138-7 Dfl 18,50
- 129 J M J DOOMERNIK, Turkse moskeeën en maatschappelijke participatie; de institutionalisering van de Turkse islam in Nederland en de Duitse Bondsrepubliek -- Amsterdam 1991: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 200 pp, 30 figs, 1 tab. ISBN 90-6809-139-5 Dfl 32,50
- 130 M DE SMIDT, A GRANBERG & E WEVER (Eds), Regional development strategies and territorial production complexes; a Dutch-USSR perspective -- Amsterdam 1991: Knag. 216 pp, 25 figs., 40 tabs. ISBN 90-6809-140-9 Dfl 29,50
- 131 P MISDORP, Centrale begrippen in de sociale geografie; een conceptuele analyse van Engelstalige leerboeken -- Amsterdam 1991: Iparto/Knag. 316 pp, 39 figs, 51 tabs. ISBN 90-6809-141-7 Dfl 39,50
- 132 M DE SMIDT & E WEVER (Eds), Complexes, formations and networks -- Utrecht/Nijmegen 1991: Knag/Geografisch Instituut Rijksuniversiteit Utrecht/ Faculteit Beelidswetenschappen Katholieke Universiteit Nijmegen. 150 pp, 32 figs, 17 tabs. ISBN 90-6809-142-5 Dfl 27,00
- 133 I I Y CASTEL, Late Holocene eolian drift sands in Drenthe (The Netherlands) -- Amsterdam/Utrecht 1991: Knag/Geografisch Instituut Rijksuniversiteit Utrecht. 162 pp, 49 figs, 15 tabs. ISBN 90-6809-143-3 Dfl 29,50
- 134 J C BORCHERT & M DE KRUYF, Bevolkingsgroei ter wille van het voorzieningsniveau? De betekenis van veranderingen in bevolkingsomvang voor het voorzieningsniveau, uitgewerkt voor de gewesten Zuid-Kennemerland en de Gooi- en Vechtstreek -- Utrecht 1991: Knag/Faculteit Ruimtelijke Wetenschappen Universiteit Utrecht. 78 pp, 4 figs, 15 tabs. ISBN 90-6809-144-1 Dfl 18,00
- 135 R VAN DER VAART, Educatief ontwerpen met geografie; een studie betreffende de structurering van geografische kennis voor educatieve doeleinden -- Utrecht 1991: Knag/Faculteit Ruimtelijke Wetenschappen Rijksuniversiteit Utrecht. 256 pp, 84 figs. ISBN 90-6809-145-X Dfl 36,00

- 136 P VAN LINDERT, Huisvestigingsstrategieën van lage-inkomensgroepen in La Paz -- Utrecht 1991: Knag/Faculteit Ruimtelijke Wetenschappen Universiteit Utrecht. 320 pp, 34 figs, 26 tabs. ISBN 90-6809-146-8 Dfl 34,50
- 137 J M M VAN AMERSFOORT & H KNIPPENBERG (Eds), States and nations: The rebirth of the 'nationalities question' in Europe -- Utrecht/Amsterdam 1991: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 198 pp, 20 figs, 12 tabs. ISBN 90-6809-147-6 Dfl 29,50
- 138 P VAN TEEFFEL, Dienstcentra en rurale ontwikkeling: een onderzoek naar het aanbod en gebruik van overheidsdiensten in Mali, Afrika -- Utrecht 1992: Knag/Faculteit Ruimtelijke Wetenschappen Universiteit Utrecht. 256 pp, 70 figs, 40 tabs. ISBN 90-6809-148-4 Dfl 34,50
- 139 T H M VAN DER LOOP, Industrial dynamics and fragmented labour markets. Construction firms and labourers in India -- Utrecht/Amsterdam 1992: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. 350 pp, 30 figs, 71 tabs. ISBN 90-6809-149-2 Dfl 49,50
- 140 H VAN DER WUSTEN (Ed), The urban political arena; Geographics of public administration -- Utrecht/Amsterdam 1992: Knag/Instituut voor Sociale Geografie Universiteit van Amsterdam. ca 190 pp. ISBN 90-6809-150-6 Dfl 35,00
- 141 B D HOEKSTRA, Informatienetwerken rondom bedrijven; de bibliotheek als informatieleverancier voor het bedrijfsleven -- Utrecht/Groningen 1992: Knag/Faculteit Ruimtelijke Wetenschappen Rijksuniversiteit Groningen. ca 100 pp. ISBN 90-6809-151-4 Dfl 25,00
- 142 H SCHRETTENBRUNNER & J VAN WESTPHAL (Eds), Empirical research and geography teaching -- Utrecht/Amsterdam 1992: Knag/Centrum voor Educatieve Geografie Vrije Universiteit Amsterdam. 200 pp. ISBN 90-6809-152-2 Dfl 30,00

Publications of this series can be ordered from KNAG/Netherlands Geographical Studies, P.O. Box 80123, 3508 TC Utrecht, The Netherlands (Fax +31 30 535523). Prices include packing and postage by surface mail. Orders should be prepaid, with cheques made payable to "Netherlands Geographical Studies". Please ensure that all banking charges are prepaid. Alternatively, American Express, Eurocard, Access, MasterCard, BankAmericard and Visa credit cards are accepted (quote card number and expiry date with your signed order).

ISSN 0169-4839

BEST COPY AVAILABLE

198